

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus.p2n model

Run on: January 17, 2004, 01:15:10 ; Search time 10931 Seconds

(without alignments)  
5344.337 Million cell updates/sec

Title: US-10-020-095-4

Sequence: 1 MGGPPLTFAHLICVCTAAL.....HSSVIFFCFLYFMEIWL 1428

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 2888711 seqs, 20454813386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Command line parameters:

-MODEL=frame+.p2n.model -DEV=rlp  
-Q=/cgn2.1/USPTO.spool.p/US10020095/runat\_15012004\_153822\_18707/app\_query.fasta\_1.1607  
-DB=GenEmbl -QFMT=fastap -SUFFIX=rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0  
-UNITS=b1bs -START=1 -END=-1 -MATRIX=blomsum62 -TRANS=human40.cdi -LIST=45  
-DOCCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000  
-USER=US10020095.@CCN\_1\_1\_7794.@runat\_15012004\_153822\_18707 -NCPU=6 -ICPU=3  
-NO\_MMAPP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

1: GenEmbl.\*  
2: gb\_da.\*  
3: gb\_hcg.\*  
4: gb\_in.\*  
5: gb\_om.\*  
6: gb\_ov.\*  
7: gb\_pat.\*  
8: gb\_ph.\*  
9: gb\_pl.\*  
10: gb\_pr.\*  
11: gb\_ro.\*  
12: gb\_scs.\*  
13: gb\_sy.\*  
14: gb\_un.\*  
15: gb\_vl.\*  
16: em\_ba.\*  
17: em\_fun.\*  
18: em\_hum.\*  
19: em\_in.\*  
20: em\_mu.\*  
21: em\_om.\*  
22: em\_or.\*  
23: em\_ov.\*  
24: em\_pac.\*  
25: em\_ph.\*  
26: em\_pl.\*  
27: em\_ro.\*  
28: em\_stg.\*  
29: em\_un.\*

## SUMMARY

SEQ ID NO: 3  
SEQ ID NO: 4  
REV. TRANS. SEQ ID NO: 4

29: em\_vl.\*  
30: em\_hcg\_hum.\*  
31: em\_hcg\_inv.\*  
32: em\_hcg\_other.\*  
33: em\_hcg\_mus.\*  
34: em\_hcg\_pln.\*  
35: em\_hcg\_rtd.\*  
36: em\_hcg\_mam.\*  
37: em\_hcg\_vrc.\*  
38: em\_hcg.\*  
39: em\_hcg\_hum.\*  
40: em\_hcg\_mus.\*  
41: em\_hcg\_other.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	7326.5	99.7	4335	6	AX537610 Sequence
2	7326.5	99.7	4761	6	AX534934 Sequence
3	7326.5	99.7	5883	9	AF410459 Homo sapi
4	7326.5	99.7	5895	6	AX534938 Sequence
5	7317.5	99.6	4335	6	AX537612 Sequence
6	7317.5	99.6	4761	6	AX534936 Sequence
7	7317.5	99.6	5895	6	AX534940 Sequence
8	5470	74.4	3535	6	AX534942 Sequence
9	5461	74.3	3535	6	AX534944 Sequence
10	5360	72.9	4735	10	AY083458 Mus muscu
11	5360	72.9	5644	10	BC052443 Mus muscu
12	4305.5	56.6	2938	9	HSB805581
13	3360	45.7	4273	6	AK095888 Homo sapi
14	1854.5	25.2	4527	6	AX054953 Sequence
15	1850	25.2	4560	6	AX054954 Sequence
16	1779.5	24.2	5603	3	CJN431688 Ciona int
17	1692	22.0	4578	3	AY122084 Drosophi
18	1692	22.0	4579	3	AJ269539 Drosophi
19	1638	22.3	4678	3	D83196 Limulus sp.
20	1542.5	21.0	5817	9	AB033109 Homo sapi
21	1542.5	21.0	5852	9	AY101765 Homo sapi
22	1491	20.3	4545	10	GPIRMA
23	1481.5	20.2	5210	9	HSB803446
24	1463.5	19.9	4598	3	DME269541
25	1460	19.9	4220	3	DME269538
26	1459	19.9	4944	3	AF538967
27	1455	19.8	4636	10	RAT113A
28	1441	19.6	4577	6	AR241630
29	1440	19.6	4425	6	AX054943 Sequence
30	1440	19.6	4577	6	AX146876 Sequence
31	1440	19.6	4577	6	AX409589 Sequence
32	1440	19.6	4577	9	HMA2M
33	1440	19.6	4617	9	BC040071 Homo sapi
34	1440	19.6	4809	6	AX281769 Sequence
35	1437	19.6	4801	5	HSB804792
36	1435.5	19.5	4841	5	LARA2M
37	1434	19.5	4569	6	A21185
38	1432.5	19.5	4622	10	AY185125 Mus muscu
39	1428.5	19.4	4637	3	AB050668 Branchios
40	1425.5	19.4	4650	3	DME269540
41	1425.5	19.4	4913	3	AY069050 Drosophi
42	1422	19.4	4488	6	AX497234 Sequence
43	1422	19.4	4488	6	AX574598 Sequence
44	1421	19.3	4599	6	A21187
45	1420	19.3	4672	10	BC051037 Mus muscu

RESULT 1

## ALIGNMENTS

AX537610  
 LOCUS AX537610 4335 bp DNA linear PAT 23-NOV-2002  
 DEFINITION Sequence 1 from Patent WO02070738.  
 ACCESSION AX537610  
 VERSION AX537610.1 GI:25269531  
 KEYWORDS  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 REFERENCE 1  
 AUTHORS Schuh, A. and Ouehland, W.  
 TITLE Diagnosis and treatment of blood disorders.  
 JOURNAL Patent: WO 02070738-A 1 12-SEP-2002;  
 Schuh, Andre (CA)  
 FEATURES  
 SOURCE Location/Qualifiers  
 1. 4335  
 /organism="Homo sapiens"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9606"  
 1. .>4335  
 /note="unamed protein product"  
 /codon\_start=1  
 /protein\_id="CAD57496.1"  
 /db\_xref="GI:25269532"  
 /translation="MQGPPLLTAAHLVCCTAALAVAPGRFLVTAPGIRPGNVTI  
 GVELLEHCPGQVTKAELLTKASNLTVSLEAEVPEKSGPKTLTLPSPNSADEY  
 ELRYLGRTODEILFNSNTRLSFETKRISVPIQDPAKYKPOEYKPRITVLSPDKFY  
 KTSNLILKDKPSNLIOQWLSQSDIGVISTPOLSHPILGDMSTIOVQNDQYYS  
 FOVEYALPKREVTLQTPLYCSMNSKHLNGITAKYVYKVKADVLTPLPESFWK  
 KKNITTKFKINGSANFSNDEKMNWDSNGLSEYLDSSPGVELITVTSVTGI  
 SRNVSTVFFKQHDYIIIEFDYTVLKPSLFTAVKTVADGNQLLEBRNNVVT  
 VTORNYTEYMGSGNSGOKMEAVOKINVTYPOSGTKPIEPILEDSELDLKYFLGS  
 KSSMAVSLSPKSPKTYIOLKTRDENIKVSGPRELAVSGKRLKEISYMTVSGQIYA  
 VGNKSTMFSLTPENSNTPKACVYVYIIEDEGLISVLKIPQIVFRKNIKLYMSKY  
 KAEPSERKSLISVQPDSTVIGIVAVKSVNLMAASNDITMENVHLELYNTGYLYG  
 MFNMSFAVFOEGCQMLVLDANLTQDYIDGVYDNAEYARFMEENEGHIVIHPSLGS  
 SPVHKHPETWILDTNMGYRIYOEFEVYVPSDITSMVANGFVLSDELGLTTPY  
 ELQAFQPEPFIPLNLPYSVIRGESEFALETIIPNLYKADNEVYIIKSDKDIIMTSB  
 INATGHOCTILVPEDEATVLPRIPTHLEGIPIYVVALSPYSALOMITLYKASBI  
 EKYSQSLIDLTNNRLOSTIKTISFSPRPPTVTSERVOITAGVLAGSINGLASL  
 IRMPYGGCEQMINIFAPVITYLDLYTKKQITDNLKRAKLSFMKQYQRELLQREB  
 SFSAFGNDPESGTMLSAFVLRCLFLEADPYIDQNLHRTYMLKGHQSNGFMDP  
 GSVHSELOGKNSPVTLAYIVTSLGKYKQPNIDVOSIHLSEFSGISDNT  
 LALITVALSSVSGPKAKALNMLTFWABOEGQMFWSSESLSDSQPSLIDEVA  
 YALISHILOPOTSGIPIMRLTSROBSIGSPASTOPTVALKALSFEALIMTERPN  
 IOVTVTGSPSSPVKFLIDTNRLLDTAEALVQPAVINISANGEPATICOANVTN  
 VKAGSSRRRRRSIQNOEAFDLDAVKENKDLNHDNLNCTSFSGPERSGALMEVNL  
 LSGPMVSEALISSETVAKVEDHKLNLIDLSVNETQPCVNI PAVENFVSNTODS  
 VSIYDYEPREAOAVRSYNSVKLSSCDLSDVQCRCEBDASGSHHSVIFPCPK  
 LLYFMELML"  
 BASE COUNT 1296 a 837 c 922 g 1280 t  
 ORIGIN  
 Alignment Scores:  
 Pred. No.: 0 Length: 4335  
 Score: 7326.50 Matches: 1427  
 Percent Similarity: 98.82% Conservative: 1  
 Best Local Similarity: 98.75% Mismatches: 0  
 Query Match: 99.71% Indels: 17  
 Gaps: 1  
 DB: 6  
 US-10-020-095-4 (1-1428) x AX537610 (1-4335)

DB 121 AATGACTATTGGGGGAGGAGCTTCGAAACATCGCCCTTCAAGAGTGACTGGAAGGCC 180  
 QY 61 GlnleuleuylthrlaseraenleuthrvalserValleuclunlagluclyValPhe 80  
 DB 181 GAGCTGCTCAAGACAGATCAAACTCATCTGTCTCTGCTGGAGAGAGAGGCTTTT 240  
 QY 81 GlulysglYserPheylsThrleuthrleuProserleuProleuanseralaaPglu 100  
 DB 241 GAAAAGGCTCTTTTAAGACATTACTTCCTTCATCACTACCTCTGAACAGTGAGATGAG 300  
 QY 101 IleTyGluLeuLysValThrGlyArgThrGlnAspGluIleuPheSerAsnSerThr 120  
 DB 301 ATTATGAGCTACGTGAACCGGACGATCCAGATGAGATTTTATCTCTTAATGATACC 360  
 QY 121 ArgLeuSerPheGlnThrLysArgIleSerValPheIleGlnThrAspLysAlaLeuTy 140  
 DB 361 CGCTTATCATTTAGACCAAGAGAAATATCTGTCTTCAATCAACAGACAGGCTTATAC 420  
 QY 141 LysProLysGlnLysValLysPheArgIleValThrLeuPheSerAspPheLysProTy 160  
 DB 421 AAGCCAAAGCAAGATGAAGTTCCCATTTGATCACTCTTCAGATTTTAAAGCCTTAC 480  
 QY 161 LysThrSerLeuAsnIleleuIleLysAspProLysSerAsnleuIleGlnGlnThrPleu 180  
 DB 481 AAAACCTCTTAAACATTTCTCATTAAGAGACCCCAATTCAAATTTGATTCACACAGTGTTG 540  
 QY 181 SerGlnGlnSerAspLeuGlyValIleSerLysThrPheGlnLeuSerSerHisPcolle 200  
 DB 541 TCACCAACAAAGATCTTGAGCATTTCCAAATCTTTCAGATCTTCCATCCCAATA 600  
 QY 201 LeuGlyAspTrpSerIleGlnValGlnValAsnAspGlnThrTyTyTyGlnSerPheGln 220  
 DB 601 CTGGTGACCTGGCTCATTCAGTTCAAGTGAAGAAAGACCAACATATATCATCATTTGAC 660  
 QY 221 ValSerGlnTyValleuProLysPheGluValThrLeuGlnThrProLeuTyTySer 240  
 DB 661 GTTTCGAATATGATTAATCCAAATTTGAAGTACCTTGACAGACACATTAATGTGTTCT 720  
 QY 241 MetAsnSerLysIleleuAsnGlyThrIleThraLysTyThrTyTyGlyLysProVal 260  
 DB 721 ATAATATCTTAAGCATTTAATGATGATCATCACGCAAGATATACATATGGAGCCAGTG 780  
 QY 261 LysGlyAspValThrleuthrPheleuProLeuSerPheTrpGlyLysLysAsnIle 280  
 DB 781 AAGGAGACTTAACGCTTACATTTTACCTTATCTTTGGGAGAAAGAAATATTT 840  
 QY 281 ThrLysThrPheLysIleleuGlySerAlaAsnPheSerPheAsnAspGluMetLys 300  
 DB 841 ACAAAAACATTTAAGATTAATGATCTGCNAACCTCTCTTTAATGATGAGAGATGAAA 900  
 QY 301 AsnValMetAsnSerSerSerAsnGlyLeuSerGluTyLeuAspLeuSerSerProGlyPro 320  
 DB 901 AATGTATAGATTTCTTCAATGATGATCTTCAATGATCTTCAATGATCTTCCCTGACCA 960  
 QY 321 ValGlnIleleuthrThrValThrGlnSerValThrGlyIleSerArgAsnValSerThr 340  
 DB 961 GTTGAATTTTAAACACAGTGAAGAAATCATGTTACAGATTTTCAAGAAATGTAAGCACT 1020  
 QY 341 AsnValPhePheLysGlnHisAspTyIleIleGluPhePheAspTyThrThrValleu 360  
 DB 1021 AATGTGTTCTTCAAGACATGATTAATCATCATGATGATGATTTTATATATCACTGCTTG 1080  
 QY 361 LysProSerLeuAsnPheThrAlaThrValLysValThrArgAlaAspGlyAsnGlnLeu 380  
 DB 1081 AAGCCATCTCTCAATCTCAAGCACGACATGAGAGTAACCTGTGCTATGAGCAACCACTG 1140  
 QY 381 ThrLeuGlnLysArgAsnAsnValIleIleThrValThrGlnLysArgAsnTyThrGln 400  
 DB 1141 ACTCTTGAGAAAGAAAGAAATATATGATCATACAGTGAACAGAGAAATCTATCATCTGAG 1200  
 QY 401 TyTrpSerGlySerAsnSerGlyAsnGlnLysMetGluAlaValGlnLysIleAsnTy 420



[illegible]

QY	781	LYSVALIIEIIEGLIYSSERASPLYPHEASPILEUEMETTRISERSEGLIILASN	800
Db	2341	AAGSTAATCATTTGGAAAAGTACAAATTTGATATTTCTAATGACTTCMAATGAATATAAT	24000
QY	801	ALATHRGLYHISGLINGINTHRLEULEVALPROSEGLUASPGIYALATHVALLEUPHE	820
Db	2401	GCCACAGGCCACCAGCAGACCCCTTGCTGTCCAGTAGAGATGGGGCAACTGTTCTTTT	24600
QY	821	PROILEARGPROTHRHILEUPLYGIIILAEPROILETHRVALILEUSERPROTHR	840
Db	2461	CCCAATCAGGCCCAACACATCTGGGAAATTCCTATACAGACAGCACAGCTCTTTCACCCACT	25200
QY	841	ALASERASPALAVALTHRGLMETILELEUVALIYSLAGIYLIIEGLIYSSERTYR	860
Db	2521	GCTTCGATGCTGTCACCCAGATGATTTTGTAAAGCTGAAGAAATAGAAAATCATAT	25800
QY	861	SERGLINSERIELEULEUASPLEUTHRASPANAIRGLEUGINSETRHLEUYSTRLEU	880
Db	2581	TCACAAATCCATCTTATTAGACTTGACACGACAAATGGGCTACAGAGTACCCTGAAAACCTTGG	26400
QY	881	SERPHESERPHERPPOPROANTHRVALITHRGLSERGLUARGVALIILETHRALILE	900
Db	2641	AGTTTCATTTCTCTCTTAATACGTACGTGCGAGTGAAGAAGTTCCAGATCACTGCAATT	27000
QY	901	GLIYASPVALLIENGLYPROSERILEASNGIYLEUALASERLEUIEARGMETPROTYRGLY	920
Db	2701	GGAGATGTTCTTGSGTCTTCATCAATGGCTTAGCTCATTTGATTCGGATGCTTATATGC	27600
QY	921	CYASGLYGLUGINAMMETIIEASPHALAPROASNIILETYRIILEUASPTYRLEUTHR	940
Db	2761	TGTGGSTAAACGAAACATGATTAATTTGGTCCCAATATTTTACATTTTGGATTATCTGACT	28200
QY	941	LYSELYSLYSGINLEUTHRASPANILEUPLYGIIYLYSALALEUSERPHEMETARGINGLY	960
Db	2821	AAAAAGAAACACCTGACAGATTAATTGAAAGAAAACCTCTTCAATTATAGCCAAAGCT	28800
QY	961	TYRGINARGIULEULEUTHRYGINARGIUNASPGIYSERPHESESRALAPHEGLIYASNTRYR	980
Db	2881	TACACAGAGAGAACTTCTCTATCAGAGGGAGATGGCTCTTCACTGCTTTGGGAAATTAT	29400
QY	981	ASPPROSEERGLYSERTHTRPLSESRALAPHEVALLEUARGYSPHELEUGIUNIASP	10000
Db	2941	GACCTTCTGGGACACTTGCTGTGCAGCTTTGTATTAAAGATTTTCTTGAAGCCGAT	30000
QY	1001	PROTYRIIEASPILEASPIINASPVALLIENHISARGTHRYTRHTRPLIYSGIYHIS	10200
Db	3001	CTTACATAGATATATGATACGAATGTGTTACAGAAACATACACTTGSGCTTAAAGACAT	30600
QY	1021	GLIYLYSERASNGIYGLIUPHETRPASPARGIYARGVALIIEHISSEERGLIULEGINGLY	10400
Db	3061	CAGAAATCCAAACGTGTAATTTTGGATCCAGGAAGATGATTCATAGCTGAGCTTCAAGCT	31200
QY	1041	GLIYASNLYSERPROVALITHRLEUTHRALTYRIILEVALITHRSELEULEUGIYTYRARG	10600
Db	3121	GGCATTAAGATCCAGTAACCTTACAGCCTATATTTGTAACCTTCTCTCGGAGATATAGA	31800
QY	1061	LYSTYRGLINPROASNIIEASPVALLINGIUSERIIEHISPEHEUGIUSERGIUPHESER	10800
Db	3181	AAGATACAGCCTTAACATTTGATGTGCAGAGAGCTATTCATTTTGGAGTCTGAATTCAGT	32400
QY	1081	ARGGLIYIIESEERASPAANTYTRTHRLEUALALEUILEHTRYRALALEUSERSEVALIGLY	11000
Db	3241	AGAGAAATTTCCAGCAATTAACCTIYAGCCCTTAATTAATTATGCAATTTGCATCAGTGGGG	33000
QY	1101	SERPROLYSALIALYSGIUALALEUASNMETLEUTHTRIPARGIAGIUGINGIUGIYGLY	11200
Db	3301	AGTCTTAAGGAAAGGAAGCTTTGAATATGTGACTTGGAGAGCAAGAACAAAGAGGTGGC	33600
QY	1121	METGLINHETRPVALISERSEGLIUSERIYLYLEUSERASPSETRPGINPROARGSERLEU	11400
Db	3361	ATGCAATTTCTGGGGTGCATAGAGTCCAAACTTTCATCTTCGGACGACCAAGCTCTCCGTG	34200

```

Qy 1141 AsplleGlValAlaAlaTyraAlaLeuSerHisPheLeuGlnPheGlnThrSerGlu 1160
Db 3421 GATATGAAAGTTGACAGCTTATGACAGCTGCTCACACCTTCAATTTCAACATTTGAG 3480
Qy 1161 GlyIleProIleMetArgTrpLeuSerArgGlnArgGlnSerLeuGlyPheAlaSer 1180
Db 3481 GGAATCCCAATATGAGGTGGCTAAGCAGGCAAGAAATACCTGGGGTGGTTCATCT 3540
Qy 1181 ThrGlnAspThrThrValAlaLeuValAlaLeuSerGluPheAlaAlaLeuMetAspThr 1200
Db 3541 ACTCAGATACCACTGCTGCTTAAAGCTCTGCTGTAATTTGACAGCCCTTAATGAAATCA 3600
Qy 1201 GluArgThrAsnIleGlnValThrValThrGlyProSerSerProSerPro----- 1217
Db 3601 GAAAGGACCAATATCCAAAGTACCGGTGACGGGGCTGAGCTCACCAAGCTCTGTAAGTTT 3660
Qy 1218 -----LeuAlaValAlaGlnPro 1223
Db 3661 CTGATTGACACACACACCGCTTACTCTTCAAGACAGAGCTTGTGCTGAGTACAGCCA 3720
Qy 1224 MetAlaValAlaSerAlaAsnGlyPheGlyPheAlaIleCysGlnLeuAsnValVal 1243
Db 3721 ATGGCAGTTAATATTTCCGCAAAATGTTTGGATTGCTATTGTCACCTCAATGTTGTA 3780
Qy 1244 TyrAsnValLysAlaSerGlySerSerArgArgArgSerIleGlnAsnGlnGluAla 1263
Db 3781 TATATATGAGAGGCTTCTGGGCTTCTTGAAGACGAGATCTATCCAAATCAAGAGCC 3840
Qy 1264 PheAspLeuAspValAlaValLysGlnAsnLysAspAspLeuAsnHisValAspLeuAsn 1283
Db 3841 TTTGATTGATGATGTTGCTGTAAGAAATGAAATGATGATCAATCAATGATGATTTGAT 3900
Qy 1284 ValCysThrSerPheSerGlyProGlyValArgSerGlyMetAlaLeuMetGluValAlaLeu 1303
Db 3901 GTGCTACACACTTTTCCGGCCCGGGTAGAGTGGCATGCTCTTATGAGAAATTAACCTTA 3960
Qy 1304 LeuSerGlyPheMetValProSerGluAlaIleSerLeuSerGluThrValLysVal 1323
Db 3961 TTAAGTGCTTTATGTCCTTCAAGAACATTTCTTGAAGCAGACGAGCAAGAAAGTG 4020
Qy 1324 GluTyrAspHisGlyLysLeuAsnLeuTyrLeuAspSerValAlaGlnThrGlnPheCys 1343
Db 4021 GAATATGATCATGAGAAACTCAACCTCATATTGATCTGTAATGAAACCCAGCTTTTGT 4080
Qy 1344 ValAsnIleProAlaValArgAsnPheLysValSerAsnThrGlnAspAlaSerValSer 1363
Db 4081 GTTAATATTTCTGCTGAGAACTTTAAAGTTTCAATTAACCAAGATGCTTCACTGCTCC 4140
Qy 1364 IleValAspTyrTyrGluProArgArgGlnAlaValArgSerTyrAsnSerGluValLys 1383
Db 4141 ATAGTGAATTAATCTATGAGCCAAAGACAGCGGTGAGAAATTAACAATCTGAAAGTGAAG 4200
Qy 1384 LeuSerSerCysAspLeuCysSerSerAspValGlnGlyCysAspProCysGluAspGlyAla 1403
Db 4201 CTGTCCTCCCTGTCGACCTTTCAGATGATCCAGGGGCTGCCCTTGTGAGAGTAGAGCT 4260
Qy 1404 SerGlySerHisHisSerSerValIlePheIlePheCysPheLysLeuLeuTyrPhe 1423
Db 4261 TCAGGCTCCATCACTCACTTTCAGTCAATTTATTTTCTGTTCAAGCTTCTGACTTT 4320
Qy 1424 MetGluLeuTrpLeu 1428
Db 4321 ATGGAACCTTGGCTG 4335

```

```

REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
source
1. 4761
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
113. 4450
/DB_xref="GI:25261475"
/note="unnamed protein product"
/codon_start=1
/protein_id="CAD57231.1"
/db_xref="GI:25261475"
/translation="MGGPPLTRAAHLCTCTAALAAVAPGPRVLTAPGIIIRRGANT
I
GVELLEHCPQVTKALELKTASNLTVSLEAGVEKESFKLTLPSPLENSADILY
ELKAVTRQDEILFNSSTRLSPEKRI SVFIQTDKALYRKQEKFRIVTLPSDFKPY
KTSINLIDKPKSNLIDQMLSOQSDIGVSKTFOLSHSPILGMSIQOVNDQTVYS
FOYSEVVLPRFEVTLQTPLYCSMNSKHLNGITPAKTYGKPGVDVTLFPLSPFGK
KNRTTTPKINSANSFNDDEMKVMDSNGLSEYLDLSSPGVPLITLVESVYGI
SRNVSTNVFPKQHDYIIFPFDYTVLKPELNFATYKTRADGNQTLTERNNVIT
VTQRNTERVSGNSNGQKMEAVQKINVTVPQSGFKIAPFLIDESSLEQLKAYFLGS
KSSMAVHSLFKSPSKTYIOLKTRDENI KVGSPPELVVSGNKRILKEISYVWSRGVLV
VAGKONSTMSLFLPENSTPKACVYVYIEDDGEIISDYKIPOLVFKKIKLYMSKV
KAEPSKVLARISVTPQDSIVGVADVCKVNIIMENVAHELELYNMGVIG
MPNNSPRAVROEGVLTLPDANLTKVIDYVDMNAEAEKRFMEBNEBHIYDIDHPSLGS
SPHVRKHPETWLTQDNGVRIYQEFETVPDSTISWATGVVISDGLGLTTPV
ELQAFQFPIFLNLPYSVIRGESEFLEIITFNLKATAEKVILEKDFILMTNE
INATGQOULVLESEDAVLPPIRPTHGELPIYVALSPDSADVMTOMILVAKGI
EKYSOSILDLIDNRLQSTLTKLSPSPNPTVGSERVOITAGVLSPLASGL
IMPVCGSGROMNFPANITVYLLDKKKOQDNMKERKALSPRQGVGVELLYORSD
SEBAFNPYPSSTTSLSAVLACFLEADYIDIDQVLRITTYTLKGRKSNPEPDP
GRVHSELQGNKSPVLTAYITVSLGRTKQPNIDVQESLHFLSEBSRGLSDYNT
LALITVLSVSGPKAKELNMTWBAEBOGQOFVSESKLSDSMPORSIDI EYAA
YALSHRLOFQISEGIPIMRWLSRORNSLGGFASIDDTVALALSEFALMNTERTN
IOVTVGSPSPVKEFLIDTHNRLLQTLAEVAVOPMAVINSANGFPAICOLAVYN
VWSSGSRRRRSIONOEAPDIDVAVKENDIDNHNIDVCTSPSGGSRSGMALMEVNL
LSGFNPVSEALISLSETVKYVEYDHGKLNLYDSVNETQCVNIPAYNRKVSNTQDAS
VSLVDYEPFRQAVRSYNSSEVKLSSCDLCSDVQCKRCEGDGASHHSVIFPCPK
LIFYMELMT"
BASE COUNT 1420 a 912 c 1016 g 1413 t
ORIGIN
Alignment Scores:
Pred. No.: 0 Length: 4761
Score: 7326.50 Matches: 1427
Percent Similarity: 98.82% Conservative: 1
Best Local Similarity: 98.75% Mismatches: 0
Query Match: 99.71% Indels: 17
DB: 6 Gaps: 1
US-10-020-095-4 (1-1428) x AX534934 (1-4761)
Qy 1 MetGluGlyProProLeuLeuThrAlaAlaLeuLeuCysValCysThrAlaAlaLeu 20
Db 113 ATGAGGAGCCACCGCTCTGACCGCGCCACCTCTCTGCTGTCAGTACCGCGCTG 172
Qy 21 AlaValAlaProGlyProArgPheLeuValThrAlaProGlyIleIleArgProGly 40
Db 173 GCCGTGGCTCCGGGCTGGCTTCTGTTGTCAGACGCCCAAGGATCATCAGCGCGGAGA 232
Qy 41 AsnValThrIleGlyValGluLeuLeuGlnHisCysProSerSerGlnValThrValLysAla 60
Db 233 AATGTACTATTTGGGTGGAGCTTCTGGAACACATGCTTCAAGAGTGAAGGCG 292
Qy 61 GluLeuLeuLysThrAlaSerAsnLeuThrValSerValLeuGlnLysIleGluGlyValPhe 80
Db 293 GAGCTGCTCAAGACACATCAACCTCACTGCTGCTGCTGAGACGACAGAGAGTCTTT 352
Qy 81 GluLysGlySerPheLysThrLeuThrLeuProSerLeuProLeuAsnSerAlaAspGlu 100

```

Db 353 GAAAAAGGCTCTTTAAGACACTTACTCTTCATCTCACTCTGAAACAGTGCAGATGAG 412  
Qy 11eTgIuLeuAryValThrgIyArGthGlnAspGluIleuPheSerAsnSerThr 120  
Db 413 ATTTAAGAGCTACGTGTAAACCGAGCTAACCGAGATGAGATTATTTCTTAATGAGACC 472  
Qy 121 ArgLeuSerPheGluThrIySarGIIeSerValPheIIeGlnThrAspLysAlaLeuTy 140  
Db 473 CGCTATCATTTGAGACCAAGAGAAATATGCTTCTTCAATTCAAACAGACAAAGGCTTATAC 532  
Qy 141 LysProLysGlnGluValLysPheArgIIeValIleuPheSerAspPheLysProTy 160  
Db 533 AACCCAAAGCAAAAGTGAAGTTCCATGTTACACTCTTCACAGATTTTAAAGCCCTTAC 592  
Qy 161 LysThrSerLeuAsnIleLeuIleLysAspProLysSerAsnLeuIleGlnGlnTrpLeu 180  
Db 593 AAAACCTCTTTAAACATTCCTCATTAAGACCCCAAAATCAATTTGATCCAAACGTGGTGG 652  
Qy 181 SerGlnGlnSerAspLeuGlyValIIeSerLysThrPheGlnLeuSerSerHisProIle 200  
Db 653 TCACAAACAAGTATCTTGAGTCATTTCCAAAACCTTTGAGCTATCTTCCCATCAAT 712  
Qy 201 LeuGlyAspTrpSerIIeGlnValGlnValAsnAspGlnThrTyTyTyGlnSerPheGln 220  
Db 713 CTGGGAGCTGGCTATTCAGATTCAAGTGAAGTGAACAGACATATTCATCATCTTACAG 772  
Qy 221 ValSerGluTyValLeuProLysPheGluValIleuGlnThrProLeuTyTySer 240  
Db 773 GTTTCGAATATGATATTAACCAAAATTTGAAGTATGATCTTGACACACCATTAATTTGTTCT 832  
Qy 241 MetAsnSerLysHisLeuAsnGlyThrIIeThrAlaLysTyThrTyTyGlyLysProVal 260  
Db 833 ATGAATTCCTAAGCATTTAATATGTTACCAACCGCAAAAGTATCAATGGAAGCCAGTG 892  
Qy 261 LysGlyAspValThrLeuThrPheLeuProLeuSerPheTrpGlyLysLysLysAsnIle 280  
Db 893 AAAGGAGACGTAACGCTTACATTTTACCTTATCTTTGGGGAAGAAATAATTT 952  
Qy 281 ThrLysThrPheLysIleAsnGlySerAlaAsnPheSerPheAsnAspGluIleLys 300  
Db 953 ACAAATAATTTAAGTAATGAATGATCGCAAACTTCTTTAATATGAAGAGTGAATA 1012  
Qy 301 AsnValMetAspSerSerAsnGlyLeuSerGluTyLeuAspLeuSerSerProGlyPro 320  
Db 1013 AATGTATATGATTTCTTAATGAGACTTTCTGAATACCTGATTAATCTTCCCTGAGACA 1072  
Qy 321 ValGluIleLeuThrThrValThrgIuSerValThrGlyIIeSerArgAsnValSerThr 340  
Db 1073 GTAGAAATTTTAAACACAGTGAAGATCAGTTACAGGTATTTCAAGAAATGTAAACACT 1132  
Qy 341 AsnValPhePheLysGlnHisAspTyTrIIeGlnPhePheAspTyThrThrValLeu 360  
Db 1133 AATGTGTCTTCAAGACAAATGATTAATCATTCATTTGATTTTGAATTAATCTGTCTTG 1192  
Qy 361 LysProSerLeuAsnPheThrAlaThrValLysValThrArgAlaAspGlyAsnGlnLeu 380  
Db 1193 AAGCCATCTCTCAACTTCACAGCACTGTGAAGGTAACTCGTCTAATGCAACCAACTG 1252  
Qy 381 ThrLeuGluGluArgArgAsnAsnValValIIeThrValThrGlnArgAsnTyThrgIu 400  
Db 1253 ACTCTGAGAGAAAGAAATATATGATCATTAACAGTGACACAGAGAAACTATATCTGAG 1312  
Qy 401 TyTrpSerGlySerAsnSerGlyAsnGlnLysMetGluAlaValGlnLysIleAsnTy 420  
Db 1313 TACTGAGGAGACTTAAAGTGAATCAAGAAATCAAGAAATCTGTTCAAGAAATTAATTTAT 1372  
Qy 421 ThrValProGlnSerGlyThrPheLysIleGluPheProIleuGlnAspSerSerGlu 440  
Db 1373 ACTGTCCCAAGTGAAGCTTTAATGATTTGAATTCCAATCTTGAGAGATTCACATGAG 1432  
Qy 441 LeuGlnLeuLysAlaTyPheLeuGlySerLysSerSerMetAlaValHisSerLeuPhe 460

Db 1433 CTACAGTTGAAGGCCATATTTCTTGTAAGTAAAGATGAGTGCAGTTCAATGCTGTTT 1492  
Qy 461 LysSerProSerLysThrTyTrIIeGlnLeuLysThrArgAspGluAsnIleLysValGly 480  
Db 1493 AAGTCTCTAAGTGAAGCATATCATCCAACTAAACAGAGAGAAATTAAGGGGGA 1552  
Qy 481 SerProPheGluLeuValValSerGlyAsnLysArgLeuLysGluLeuSerTyMetVal 500  
Db 1553 TCGCCTTTGAGTGGTGGTTAGTGGCAACAACGATTCGAAGAGTTAAGTAAATGTA 1612  
Qy 501 ValSerArgGlyGlnLeuValAlaValGlyLysGlnAsnSerThrMetPheSerLeuThr 520  
Db 1613 GTATCCAGGGGACAGTGGTGGCTGAGGAAAACAAAATTCACAAATGTTCTTTAACA 1672  
Qy 521 ProGluAsnSerTrpThrProLysAlaCysValIIeValTyTyTrIIeGlnAspAspGly 540  
Db 1673 CAGAAATATCTTGACTCCAAAGCCGTGTAAATGATTAATTAAGATGATGGG 1732  
Qy 541 GluIleIleSerAspValLeuLysIleProValGlnLeuValPheLysAsnLysIleLys 560  
Db 1733 GAATTTATAGAGATGTTCTAATAAATCTCTGTCAGCTTGTTTAAATAAGATTAAG 1792  
Qy 561 LeuTyTrpSerLysValLysAlaGluProSerGluLysValSerLeuArgIIeSerVal 580  
Db 1793 CTATATGAGATTAAGTGAAGTGAACCTGACCATCTGAAGAGTCTCTTATGAGATCTGTG 1852  
Qy 581 ThrGlnProAspSerIleValGlyIleValAlaValAspLysSerValAsnLeuMetAsn 600  
Db 1853 ACACAGCCTGACTCCATGTTGGAGTTGAGTGTGTCACAAAGTGAATGTATGTAAT 1912  
Qy 601 AlaSerAsnAspIleThrMetGluAsnValValHisGluLeuGluLeuTyAsnThrgly 620  
Db 1913 GCTCTAATGATATTCATGAAATGAAGATGGTCCATGAGTTGAACTTTATTAACACAGA 1972  
Qy 621 TyTyTyLeuGlyMetPheMetAsnSerPheAlaValPheGlnGluCysGlyLeuTyVal 640  
Db 1973 TATTAATTAAGCATGTTCATGAATCTTTTGAGCTTTCAGGAATGTGACTCGGGTA 2032  
Qy 641 LeuThrAspAlaAsnLeuThrLysAspTyTrIIeAspGlyValTyAspAsnAlaGluTy 660  
Db 2033 TTGACAGATGCAAACTCAAGAGATTAATATGATGTGTTATGACAAATCCAAATAT 2092  
Qy 661 AlaGluArgPheMetGluGluAsnGluGlyHisIleValAspIleHisAspPheSerLeu 680  
Db 2093 GCTGAAGGTTTATGAGAAATATGAAGCATATGTAATGATTAATCAATGACTTTCTTGG 2152  
Qy 681 GlySerSerProHisValArgLysHisPheProGluThrTrpIleTrpLeuAspThrAsn 700  
Db 2153 GGTACAGTCCACATGTCGAAAGCATTTTCCAGAGACTTGGATTGGCTAGACCAAC 2212  
Qy 701 MetGlyTyTrArgIIeTyTrGlnGluPheGluValThrValProAspSerIleHisSerTrp 720  
Db 2213 ATGGGTTCACGATTTACCAAGAAATTTGAAGTGAACCTGATTCATCACTCTCTGG 2272  
Qy 721 ValAlaThrglyPheValIIeSerGluAspLeuGlyLysGlyLeuThrThrProVal 740  
Db 2273 GTGGCTACTGGTTTGTGATCTCTGAACCTGGGCTTTGACATCAACATCACTCAAGT 2332  
Qy 741 GluLeuGlnAlaPheGlnProPhePheIlePheLeuAsnLeuProTySerValIIeArg 760  
Db 2333 GAGCTCCAAAGCTTCCAAACATTTTTCATTTTGTGAATCTTCCCTATCTGTTATCGA 2392  
Qy 761 GlyGluGluPheAlaLeuGluIleThrIlePheAsnTyTrLeuLysAspAlaThrGluVal 780  
Db 2393 GGTGAAGATTTGCTTGTGAATATCAATTAATTTTGAAGAGTCCACTGAGGTT 2452  
Qy 781 LysValIleIleGluLysSerAspLysPheAspIleLeuMetThrSerSerGluIleAsn 800  
Db 2453 AAGGTATCATTTGAAGAAATGACAAATTTGATTTCAATGACTTCAAAAGAAATTAAT 2512  
Qy 801 AlaThrglyHisGlnGlnThrLeuLeuValProSerGluAspGlyAlaThrValLeuPhe 820  
Db 2513 GCCACAGGCCACACAGACACCTTCTGTGTTCCAGTGAAGATGGGGCAACTGTTCTTTT 2572

QY 821 ProileArgProthHisLeuGlyGluIleProIleThrValIleuMetSerProThr 840  
 DB 2573 CCATACAGGCCAACATCTGGAGAAATTCCTATACAGTCAAGGCTTTTCAACCACT 2632  
 QY 841 AAserAspAlaValThrGluMetIleLeuValIleGluGlyIleGluLeuSerTyr 860  
 DB 2633 GCTTCTGATGCTGCACCCCAATGATTTTAAAGGCTGAAGAAATGAAATTCATAT 2692  
 QY 861 SerGlnSerIleLeuLeuAspLeuThrAspAsnArgLeuGlnSerThrLeuTyrThrLeu 880  
 DB 2693 TCACATATCATCTTATAGACTTACATACATAGCTACAGATACCTGAAACCTTTG 2752  
 QY 881 SerPheSerPheProProAsnThrValThrGlySerGluArgValGlnIleThrAlaIle 900  
 DB 2753 AGTTTCATTTCTCTCTAATACAGTGAAGTGAAGAGTTCAAGATCATGCAATT 2812  
 QY 901 GlyAspValIleuGlyProSerIleAsnGlyLeuAlaSerLeuIleArgMetProTyrGly 920  
 DB 2813 GGAGATGTTCTTGCTTCATCATATGCTTAGCTTCATGATTCGATGCCATTATGCC 2872  
 QY 921 CysGlyGluGlnAsnMetIleAsnPheAlaProAsnIleTyrIleLeuAspTyrLeuThr 940  
 DB 2873 TGTGTGAACAGAACATCATATAATTTGCTCAAAATATTACATTTTGATTTATCTGACT 2932  
 QY 941 LysIleGlyGlnLeuThrAspAsnLeuGlyGluValAlaLeuSerPheMetArgGlnGly 960  
 DB 2993 AAAAAGAAACAACACAGATATATTGAAAGAAAGCTCTTTCAATTTATAGGCAAGCT 2992  
 QY 961 TyrGlnArgGluLeuLeuTyrGlnArgGluAspGlySerPheSerAlaPheGlyAsnTyr 980  
 DB 2993 TACAGAGAGAACTTCTCTATCAGAGAGAGATGCTCTTCAAGTCTTTGGGAAATTTAT 3052  
 QY 981 AspProSerGlySerThrTyrLeuSerAlaPheValLeuArgCysPheLeuGluAlaAsp 1000  
 DB 3053 GACCTTCTGGAGACTTGTGTGACCTTTGTTTAAAGATGTTTCTTGAAGCCGAT 3112  
 QY 1001 ProTyrIleAspIleAspGlnAsnValIleuHisArgThrTyrThrTyrLeuGlyHis 1020  
 DB 3113 CCTTACATAGATTTGATCAGATATGTTTACACAGAACATACATCTGGCTTAAAGGACAT 3172  
 QY 1021 GlnLeuSerAsnGlyGluPheTyrAspProGlyArgValIleHisSerGlnLeuGlnGly 1040  
 DB 3173 CAGAAATCAACGGTGAATTTGGGATCCAGAAAGATGATTCATAGAACCTTCAAGGT 3232  
 QY 1041 GlyAsnLeuSerProValThrLeuThrAlaTyrIleValThrSerLeuLeuGlyTyrArg 1060  
 DB 3233 GGCATATAAAGTCCAGTACACTTACAGCTTATATGTAATTTCTCTCTGGGATATAGA 3292  
 QY 1061 LysTyrGlnProAsnIleAspValGlnLeuSerIleHisPheLeuGluSerGluPheSer 1080  
 DB 3293 AAGATACAGCTTACATTTGATGTCGAAGATCTATCCATTTTGGAGTCTGAATTCAGT 3352  
 QY 1081 ArgGlyIleSerAspAsnTyrThrLeuAlaLeuIleThrTyrAlaLeuSerSerValGly 1100  
 DB 3353 AGAGCAATTTTCAGCAATTAATACCTTACCTTAACTTATGCAATGTCATGCTGGGG 3412  
 QY 1101 SerProValAlaValGluAlaLeuAsnMetLeuThrTyrPheValGluGlnGluGly 1120  
 DB 3413 AGCTCTAAAGCGAAGAGCTTTGAATAGCTTGAAGAGCGAAGCAAGAGAGTGGC 3472  
 QY 1121 MetGlnPheTyrValSerSerGlySerLeuSerAspSerTyrGlnProArgSerLeu 1140  
 DB 3473 ATGCAATTTCTGGGTCTCATCAGAGTCCAAACTTTTCTGACTCTGGAGCCAGCTCCCTG 3532  
 QY 1141 AspIleGluValAlaAlaTyrAlaLeuLeuSerHisPheLeuGlnPheGlnThrSerGlu 1160  
 DB 3533 GATATTGAAGTTGAGCTTATGCACTGCTCACACTTCTTACATTTCAAGATTTCTGAG 3592  
 QY 1161 GlyIleProIleMetArgTyrPheLeuSerArgGlnArgAsnSerLeuGlyGlyPheAlaSer 1180  
 DB 3593 GGAATCCCAATTAATGAGGTGCTTAGAGGCAAGAAATAGCTTGGGTGTTTGCATCT 3652

QY 1181 ThrGlnAspThrThrValAlaLeuValAlaLeuSerGluPheAlaAlaLeuMetAsnThr 1200  
 DB 3653 ACTCAGGATACCACTGGGCTTTAAAGGCTCTGTCTGAATTTGCACGCCCTTAATGAATACA 3712  
 QY 1201 GluArgThrAsnIleGlnValThrValThrGlyProSerSerProSerPro----- 1217  
 DB 3713 GAAAGCAAAATATTCAGAGTACCGGGCTTAGCTCACACAGTCTCTGTAAGTTT 3772  
 QY 1218 -----LeuAlaValAlaGlnPro 1223  
 DB 3773 CTGATTGACACACAAACCGCTTACTCTTACAGACAGAGAGTCTGTGTGATACAGCCA 3832  
 QY 1224 MetAlaValAsnIleSerAlaAsnGlyPheGlyPheAlaIleCysGlnLeuAsnValVal 1243  
 DB 3833 ATGGCAGTTAAATTTTCCGCAAAATGGTTTGGATTTGCTAATTTGTCAGCTCAATGTTGA 3892  
 QY 1244 TyrAsnValIleValAserGlySerSerArgArgArgSerIleGlnAsnGlnGluAla 1263  
 DB 3893 TATTAATGMAAGCTTCTGGGCTTCTAGAGAGCAAGATCTATCCAAATCAAGAGGCC 3952  
 QY 1264 PheAspLeuAspValAlaValIleGlyGluAsnIleAspAspLeuAsnHisValAspLeuAsn 1283  
 DB 3953 TTTGATTTAGATGTGCTGTGAAGAAATTAAGATGATCTCATCATGATGATTTGAAT 4012  
 QY 1284 ValCysThrSerPheSerGlyProGlyArgSerGlyMetAlaLeuMetGluValAsnLeu 1303  
 DB 4013 GTGTGTACAAAGCTTTTCCGGCCCGGGTAGAGTGGCATGCTCTTATGGAATTTAACTTA 4072  
 QY 1304 LeuSerGlyPheMetValProSerGluAlaIleSerLeuSerGluThrValIleValVal 1323  
 DB 4073 TTAAGTGGCTTATATGGCTTCTAGAGCAATTTCTGTGACAGACAGTGAAGAAAGTG 4132  
 QY 1324 GluTyrAspHisGlyIleLeuAsnLeuTyrLeuAspSerValAsnGluThrGlnPheCys 1343  
 DB 4133 GAATATGATCATGGAAGAACTCAACCTCTATTGATTTCTGTAATGAAGAACCACTTTGT 4192  
 QY 1344 ValAsnIleProAlaValAlaArgAsnPheValSerAsnThrGlnAspAlaSerValSer 1363  
 DB 4193 GTTAATATTCCTCTGTGAGAAACTTTAAAGTTTCAATCCCAAGATGCTTCAAGTCTCC 4252  
 QY 1364 IleValAspTyrTyrGluProArgArgGlnAlaValArgSerTyrAsnSerGluValIle 1383  
 DB 4253 ATAGTGATTAATTAATGAGCCAGAGAGACAGCGGTAGAAATTAACACTTGAAGTGAAG 4312  
 QY 1384 LeuSerSerCysAspLeuCysSerAspValGlnGlyCysArgProCysGluAspGlyAla 1403  
 DB 4313 CTGTCTCTCTGTACCTTTGCAGTGAATGTCAGAGGCTGCCGTCTGTAGATGAGACT 4372  
 QY 1404 SerGlySerHisHisHisSerSerValIlePheIlePheCysPheLeuLeuTyrPhe 1423  
 DB 4373 TCAGGCTCCCATCATCACTCTCAGTCAATTTATTTCTGTTCAAGCTTCTGTACTTT 4432  
 QY 1424 MetGluLeuThrLeu 1428  
 DB 4433 ATGGAATTTGGCTG 4447  
 RESULT 3  
 AF410459 5883 bp mRNA linear PRI 02-MAR-2002  
 LOCUS Homo sapiens CD109 (CD109) mRNA, complete cde.  
 DEFINITION AF410459  
 VERSION AF410459.1 GI:19071208  
 KEYWORDS  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 REFERENCE 1 (bases 1 to 5883)  
 AUTHORS Lin, M., Sutherland, D.R., Horsfall, W., Tooty, N., Yeo, E., Nayyar, R.,  
 Wu, X.F., and Schuch, A.C.  
 TITLE Cell surface antigen CD109 is a novel member of the alpha(2)  
 macroglobulin/C3, C4, C5 family of cholesteryl-containing proteins  
 JOURNAL Blood 99 (5), 1683-1691 (2002)

MEDLINE 21849742  
PUBMED 11861284  
REFERENCE 2 (bases 1 to 5883)  
AUTHORS Lin, M., Sutherland, D. R., Horsfall, W., Totty, N., Yeo, E., Nayar, R.,  
TITLE Wu, X.-F. and Schaub, A. C.  
JOURNAL Submitted (14-AUG-2001) Medicine, University of Toronto, 1 King's  
College Circle, Room 7366, Toronto, Ontario M5S 1A8, Canada  
FEATURES  
source  
1. 5883  
Location/Qualifiers  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="K1"  
1. 5883  
/gene="CD109"  
1. 112  
/gene="CD109"  
113. 4450  
/gene="CD109"  
/note="associated with the Gov alloantigen system"  
/codon\_start=1  
/product="CD109"  
/protein\_id="AA184159.1"  
/db\_xref="GI:19071209"  
/translation="MQGPLLTAHLVCVCTAALAVAPGRPLVTAPGIIIRGNVTI  
GVELLEHCPQVTVKAEILKTASNLTVSLAEVGEKGSFKTLPLPLNADAEIY  
ELRVTRTODEILFNSSTRLSPEKRIISVFIOTDKALYKQEVKRIYTLPSDRKY  
KTNILIKDKPSNLIQWLISQSDLGIVSKTPOLSSHIIIDMSIQUVNDQTYQS  
RVSSEVLKPEVYTLQPLCYGSMKSHLNGITTAKTGKPKVAGDITFLFLSTMGK  
KNKITEFKINGSNDIIEPFDYTVLKSLNFTAVYKLVRAAGNOLTEERNNVIT  
VQNRVTEYMSNSGNQKMAVOKINTVPOSGFKIEPILDESSLOLKAFLGS  
KSMVAHSEFKSPKTYIQLKTRDENIKVSPFELVSENGRKLKELSWVYRGOLV  
VQKQSTMSLTPENSWPKACVIVYIEDDEIISDUKITYQVFKKIKLYMSKY  
KAEBSKVSRLSVTOPDSIVGIVANDSVNMAENDITMENVAHELELYVYLG  
MENSFAVQEGGLWLDANLTKOYIDVYNAEVAEEMENGGHIVDHPILGS  
SHVRKHPEPFIWLDTMNGYRIQEFETVPDSITSWATGVISEDGLATTPV  
ELQAPQPFRIPLNPSYVIGSEPALEITIPYIKDQATVKVIEKSDPFLMTSE  
INATGHOOFLVPSBDGATVLPPIPTHGELPIYVTLSPASDAVHOMILVAKGI  
EKSYOSILDLTDNRLOSTLTKTSFPFNTVTSERQITAGIDVAGPSINGLASL  
IMPTGCGEONMIFAPNVIILDYLTKKQLTDNLKELKSMRQGYORELLYQEDG  
SFSAGNVDPSGKSWLAFVLCPLFADPYIDIDNVLRITTYMLKSGEPMFPD  
GVAHSELDQGSKSPVTLTAYIVTSLGKRYOPNIDVQESIHFLPESSEKISNNT  
IALITYLASSVGSPPAKKALMLTWRAEGGMFVSESLSDSMOBRSLDIVAA  
YALLSHLOFQPSSEGIPIRMWLSRQNSIGRSQDPTVAALKASERPAALNTERTN  
IOVTVTGSSSPVAFRLIDTHNRLLQTELELAVOPMAVNISANGGFATCOLNVVN  
VKAASSRRRRSIOQEAFLDVAVAKENDIDNVDLVNCTFSGGRGNALMEVN  
LSGFVWSEPAISLSETVKVEYDHGLNLYDSVNEIOPCVNI PAVRNFKVNTODAS  
VSIIVYVEPRQAVRSYNSVEXLSCDLCSVDVQGRPCEDASGSHHSVIFIFCFK  
LLYFMEIWL"  
3'UTR  
4451..5883  
/gene="CD109"  
misc\_feature  
4748  
/gene="CD109"  
/note="alternative polyA site found on clone K1"  
BASE COUNT 1697 a 1173 c 1241 g 1772 t  
ORIGIN  
Alignment Scores:  
Pred. No.: 0 Length: 5883  
Score: 7326.50 Matches: 1427  
Percent Similarity: 98.82% Conservative: 1  
Best Local Similarity: 98.75% Mismatches: 0  
Query Match: 99.71% Indels: 17  
DB: 9 Gaps: 1  
US-10-020-095-4 (1-1428) x AP410459 (1-5883)  
QY 1 MetGInGlyYProProleuLeuThrAlaAlaHISleuLeuCyValCyGThrAlaAlaLeu 20  
Db 113 ATGCAAGGCCACCGCTCTGACCGCGCCACCTCTCTGCGTGTGACCGCGCGCTG 172

QY 21 AlaValAlaProGlyProArgPheLeuValThrAlaProGlyIleIleArgProGlyGly 40  
Db 173 GCCGTGGCTCCCGGGGCTCGGCTTCTGCTGACAGCCCGAGGATCATCAGGCCGAGAGA 232  
QY 41 AsnValThrIleGlyValGluLeuLeuGluHisCysProSerGlnValThrValValAla 60  
Db 233 AATGAGACTAATGGGGGTGAGCTTCGTGAGAACATGCCCCCTTACAGGTACTGTGAAAGCG 292  
QY 61 GluLeuLeuLysThrAlaSerAsnLeuThrValSerValLeuGluAlaGlyValAlaPhe 80  
Db 233 GAGCTGCTCAAGACAGCATCAAACTCATCTGCTGCTCGAAGCAGAAAGGACTCTT 352  
QY 81 GlyLysGlySerPheLysThrIleuThrLeuProSerIleuProLeuAsnSerAlaAspGlu 100  
Db 353 GAAAAAGGCTCTTTAAGACACTTACTTCCATCATCACTCTGACATGACATGATGAG 412  
QY 101 IleTyrgIuLeuArgValThrGlyArgThrGlnAspGluIleLeuPheSerAsnSerThr 120  
Db 413 ATTATAGAGCTAAGGTATACCGGAGCTACCGAGATGAGATTTATCTCTAATGTATCC 472  
QY 121 ArgLeuSerPheGluThrLysArgIleSerValPheIleGlnThrAspLysAlaLeuTyrr 140  
Db 473 CGCTTATCATTTAGACACCAAGAGATATCTGCTTCTTCAATCAACAGACAGGCTTATAC 532  
QY 141 LysProLysGlnGluValLysPheArgIleValThrLeuPheSerAspPheLysProTyrr 160  
Db 533 AAGCCAAAGCAAGAGATGAAAGTTGGATGCTTACCTTCACTCTGATTTTAAGCTTAC 592  
QY 161 LysThrSerLeuAsnIleLeuIleLysAspProLysSerAsnLeuIleGlnGlnThrLeu 180  
Db 593 AAACCTCTTTAAACATTTCTCATTAAGACCCCAATCAATTTGATCCAAACAGGCTTG 652  
QY 181 SerGlnGlnSerAspLeuGlyValIleSerLysThrPheGlnLeuSerSerHisProIle 200  
Db 653 TCACACAAAGTATGATCTTGAGTCACTTCCAAATTTTCAGTATCTTCCATCCAAATA 712  
QY 201 LeuGlyAspTyrSerIleGlnValGlnValAsnAspGlnThrTyrrTyrrGlnSerPheGln 220  
Db 713 CTGGTGACGTGTCATTTCAAGTTCAAGTATGATGACGACATNTATCATCATTTTCAG 772  
QY 221 ValSerGluTyrrValLeuProLysPheGluValThrLeuGlnThrProLeuTyrrCysSer 240  
Db 773 GTTTCAGATATGTATTTCCAAATTTGAAAGTCACTTTCAGACACCATATATATGTTCT 832  
QY 241 MetAsnSerLysHisLeuAsnGlyThrIleThrAlaLysTyrrTyrrGlyLysProVal 260  
Db 833 ATGAATTTTAAACATTTAATGAATGATCACTCAAGCAAGATATATGGAAGCCAGTG 892  
QY 261 LysGlyAspValThrLeuThrPheLeuProLeuSerPheTyrGlyLysLysAsnIle 280  
Db 893 AAAGAGACGTATACGCTTACATTTTACCTTTATCTTTGGGAAAGAAAGAAAATATTT 952  
QY 281 ThrLysThrPheLysIleAsnGlySerAlaAsnPheSerPheAsnAspGluIleLys 300  
Db 953 ACAAATAATTTAAGATTAATGATGATCTGCAAACTCTCTTTAATGATGAAGATGANA 1012  
QY 301 AsnValMetAspSerSerAsnGlyLeuSerGluTyrrLeuAspLeuSerSerProGlyPro 320  
Db 1013 AATGTAAATGATTTCTTCAATGAGCTTCTCAAAATCCGGAATCATCTTCCCTGAGACA 1072  
QY 321 ValGluIleLeuThrThrValThrGluSerValThrGlyIleSerArgAsnValSerThr 340  
Db 1073 GTAGAAATTTTAAACCAAGTACAGATCAAGATCAAGTATTTCAAGAAATGTAAGCACT 1132  
QY 341 AsnValPhePheLysGlnHisAspTyrrIleIleGluPhePheAspTyrrThrThrValLeu 360  
Db 1133 AATGTGTTCTTCAAGCAACATGATTAATCATGAGTTTGTGATTAATCACTACTGTTTG 1192  
QY 361 LysProSerLeuAsnPheThrAlaThrValLysValThrArgAlaAspGlyLysAsnGlnLeu 380  
Db 1193 AAGCCATTCCTCAACTTCAAGCAGCATGTGAAAGTAACTCTGCTGATGAGCAACCAACTG 1252  
QY 381 ThrLeuGluGluArgArgAsnValValIleThrValThrGlnArgAsnTyrrThrGlu 400



Db	1253	ACTCTTGAAGAAAGAAATAATGATGATCAACAGTGCACACAGAAACATATCTGAG	1312
Qy	401	TyrtPserGlySerAsnSerGlyAsnGlnLysMetGluAlaValGlnLysIleAsnTy	420
Db	1313	TACTGGAGCGGATCTTAACAGTGAATCAGAAATCGAAAGCTGTTCAAGAAATTAATTTAT	1372
Qy	421	ThyValProGlnSerGlyThrPheLysIleGluPheProIleLeuGluAspSerSerGlu	440
Db	1373	ACTGTCGCCCAAGATGGAACCTTTTAAGATGGAATCCCAATCCTGGAGGATTCACGTGAG	1432
Qy	441	LeuGlnLeuLysAlaTyrrPheLeuGlySerLysSerSerMetAlaValHisSerLeuPhe	460
Db	1433	CTAAGATTGAAGGGCTATTTCTCTTGTAAGTAAAGACATGGCACTTCATGCTGTTT	1492
Qy	461	LysSerProSerLysThrTyrrIleGlnLeuLysThrArgAspGluAsnIleLysValGly	480
Db	1493	AAGTCTCTGATGAACATACATCTCAACCTAAAAACAGATCAAAATTTAAAGGTGGGA	1552
Qy	481	SerProPheGluLeuValIleSerGlyAsnLysArgLeuLysGluLeuSerTyrrMetVal	500
Db	1553	TCGCTTTTGAAGTGGTGGTTAGTGGCAACAAACGATGAGAGATTAGCTATATGGTA	1612
Qy	501	ValSerArgGlyGlnLeuValAlaValGlyLysGlnAsnSerThrMetPheSerLeuThr	520
Db	1613	GTATCCAGGGGACAGTGGTGGCGTAGAAGAAACAAATTCACAAATGTTCTCTTAAACA	1672
Qy	521	ProGluAsnSerTyrTrpThrProLysAlaCysValIleValTyrrTyrrIleGluAspArg	540
Db	1673	CCAGAAATTTCTTGACATCCAAAGACCTGTGTATTTGTATTTATTTAAAGATGATGGG	1732
Qy	541	GluIleIleSerAspValLeuLysIleProValGlnLeuValPheLysAsnLysIleLys	560
Db	1733	GAATTTATTAAGTAGTGTCTTAATAAATTCCTGTTCACTTGTTTTTAAAAATAGATTAAAG	1792
Qy	561	LeuTyrrTrpSerLysValLysAlaGluProSerGluLysValSerLeuArgIleSerVal	580
Db	1793	CTAATATGGAGTAAAGGAAAGCTGAAACCATCTGAAAGAAAGTCTCTTGAGATCTCTGG	1852
Qy	581	ThrGlnProAspSerIleValGlyIleValAlaValAspLysSerValAsnLysMetAsn	600
Db	1853	ACACAGCCTGATCTCCATAGATGGGATTTAGCTGTGACAAAGATGTGAACTGATGAAT	1912
Qy	601	AlaSerAsnAspIleThrMetGluAsnValIleHisGluLeuGluLeuTyrrAsnThrGly	620
Db	1913	GCCTCTATGATATATTAACAATGGAAGAAATGGTCCATGAGTGGAACTTTATATACACAGA	1972
Qy	621	TyrrTyrrLeuGlyMetPheMetAsnSerPheAlaValPheGlnGluCysGlyLeuTrpVal	640
Db	1973	TATATATTAGCAATGTTCAATGAATTCCTTTTGACAGTCTTTCAAGAAATGTGACTCGGGTA	2032
Qy	641	LeuThrAspAlaAsnLeuThrLysAspTyrrIleAspGlyValTyrrAspAsnAlaGluTy	660
Db	2033	TTGACAAATGAAACCTCCACGAAGGATATATATGATGGTGTATGACAAATCCAAATAT	2092
Qy	661	AlaGluArgPheMetGluGluAsnGluGlyHisIleValAspIleHisAspPheSerLeu	680
Db	2093	GCTAGAGGTTTATGAGGAAAAATGAAGACATATTTGTAGATATCATGACTTTCTTTG	2152
Qy	681	GlySerSerProHisValArgLysHisPheProGlnThrTrpIleTrpLeuAspThrAsn	700
Db	2153	GGTAGCAGTCCACATGTCCGAAAGACATTTTCCAGACACTTGGATTGGCTAGACACAAAC	2212
Qy	701	MetGlyTyrrArgIleTyrrGlnGluPheGluValThrValProAspSerIleHisSerTrp	720
Db	2213	ATGGGTTTACAGGATTTACCAAGAAATTTGAAGTACGTACTGATTTATATCACTTCTGG	2272
Qy	721	ValAlaThrGlyPheValIleSerGluAspLeuGlyLeuGlyLeuThrThrThrProVal	740
Db	2273	GTGGCTACTGGTTTGTGATCTCTGAGAGACTGGGCTTGGACTAACAATCTCAAGTG	2332
Qy	741	GluLeuGlnAlaPheGlnProPhePheIlePheLeuAsnLeuProTyrrSerValIleArg	760

QY 1121 MetGlnPheTyrValSerSerGluSerIleuSerAspSerTyrGlnProArgSerIleu 1140  
 DB 3473 ATGCATTCCTGGGTCATCATCAGATCCAACTTTCTGACTCTGGCAGCAGCTCCCTG 3532  
 QY 1141 AspIleGluValAlaAlaTyrAlaIleuSerSerIlePheLeuGlnPheGlnThrSerGlu 1160  
 DB 3533 GATATTTGAAGTTGACGGCTATGACCTGCTCAGCTTCTTCAATTTCCAGCTTCTGAG 3592  
 QY 1161 GlyIleProIleMetArgTyrIleuSerArgIleArgSerIleuGlyIlePheAlaSer 1180  
 DB 3593 GGAATCCCAATTTAGAGTGGCTTAACAGCAGCAAGAAATAGCTTGGGTCGTTTTCATTT 3652  
 QY 1181 ThrGlnAspThrThrValAlaIleuIleuSerIleuPheAlaAlaIleuMetAsnThr 1200  
 DB 3653 ACTCAGAGATACCACTGCTGCTTTAAAGCTCTGTCTGATTTTGACGCCCTTAATGATATA 3712  
 QY 1201 GluIleGlnThrAsnIleGlnValThrValThrGlyProSerSerProSerPro----- 1217  
 DB 3713 GAAAGGACAAATATCCAGTGAACGGGAGCTTCACTCAAGTCTCTGTTAAAGTTT 3772  
 QY 1218 -----LeuAlaValAlaGlnPro 1223  
 DB 3773 CTGATTGACACACACACCGCTTACTCTTCAAGACAGACCTTGTGTGTAACGCCA 3832  
 QY 1224 MetAlaValAsnIleSerAlaAsnGlyPheGlyPheAlaIleCysGlnIleuAsnValVal 1243  
 DB 3833 ATGGCAGTTAATATTTCCGCAATGCTTTGATTTGCTATTTGTCAGCTCAATGTGTA 3892  
 QY 1244 TyrAsnValIleValSerGlySerSerArgArgArgSerIleGlnAsnGlnIleVal 1263  
 DB 3893 TATAATGTGAAGGCTTCTGGGCTTCTTCAAGACGAAAGTCTATCCAAATCAAGAGCC 3952  
 QY 1264 PheAspIleuAspValAlaValIleGlnIleuAspIleuAsnIleValIleuAsn 1283  
 DB 3953 TTTGATTTAGATGTGCTGTAAAGAAATTAAGATGATCTCAATCATGTGATTTGAAT 4012  
 QY 1284 ValCysThrSerPheSerGlyProGlyArgSerGlyMetAlaIleuMetGluValAsnIleu 1303  
 DB 4013 GTGTGTACAGCTTTTCGGGCCCGGGTGAAGATGGTGCATTTGTAATGAACTTAACCTA 4072  
 QY 1304 LeuSerGlyPheMetValProSerGluAlaIleSerIleuSerGluThrValIleVal 1323  
 DB 4073 TTAAGTGGCTTATGTGCTTCAAGACAAATTTCTCTAGCGAGACAGTGAAGAAAGTG 4132  
 QY 1324 GluTyrAspPheGlyIleuIleuAsnIleuTyrIleuAspSerValAsnGlnThrGlnPheCys 1343  
 DB 4133 GAATATGATCATGAGAACTCAACCTCTATTTAGATTTCTGTAATGAAACCCAGTTTGT 4192  
 QY 1344 ValAsnIleProAlaValArgAsnIleuValSerAsnThrGlnAspAlaSerValSer 1363  
 DB 4193 GTTAATATTTCTGCTGTGAGAACTTTAAGTTTCAATATCCAAAGATGCTTCAAGTCTC 4252  
 QY 1364 IleValAspTyrTyrGluProArgArgGlnAlaValArgSerTyrAsnSerGluValIle 1383  
 DB 4253 ATAGTGTGATTAATAGAGCCCAAGAGACAGCGGTGAGAAAGTTTCAACTCTGAAGTGAAG 4312  
 QY 1384 LeuSerSerCysAspIleuCysSerAspValGlnGlyCysArgProCysGluAspGlyAla 1403  
 DB 4313 CTGTCTCTCTGTGACCTTTGAGTGAATGTCAGAGGCTGCGCTTGTGATGATGAGCT 4372  
 QY 1404 SerGlySerPheIleIleIleSerSerValIlePheIlePheCysPheIleuIleuTyrPhe 1423  
 DB 4373 TTAGGCTCCCATCATCATCTTCTCATGATTTTATTTTCTGTTTCAAGCTTCTGTACTTT 4432  
 QY 1424 MetGluLeuTyrIleu 1428  
 DB 4433 ATGGAACCTTGGCTG 4447  
 RESULT 4  
 AX534938  
 LOCUS AX534938 5895 bp DNA linear PAT 22-NOV-2002  
 DEFINITION Sequence 5 from Patent WO02070696.

ACCESSION AX534938  
 VERSION AX534938.1 GI:25261482  
 KEYWORDS  
 SOURCE  
 ORGANISM Homo sapiens (human)  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrate; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 REFERENCE  
 AUTHORS Schuch, A. and Sutherland, R. D.  
 TITLE Cdi09 nucleic acid molecules, polypeptides and methods of use  
 JOURNAL Patent: WO 02070696-A 5 12-SEP-2002;  
 Schuch, Andre (CA); Sutherland, Robert D. (CA)  
 FEATURES  
 source  
 1..5895  
 /organism="Homo sapiens"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9606"  
 113..4450  
 /note="unnamed protein product"  
 /codon\_start=1  
 /protein\_id="CAD57233.1"  
 /db\_xref="GI:25261483"  
 /translation="MQGPPILTAALHLCVTAALAVAPGPPFLVTAGIIRPGNVTI  
 GVLELHCPQGVTVKAEILKTASMLTVSVLEAEVFEKSGFKTLTLPGLPLNSADEIY  
 ELRYVGRQDEILFNSNTRLSFEKRLISVFLQTKALYKPKQEVAFRLVTLFSPKPY  
 KTSNLNLIKDKSKNLILOMISOQDLGIVSKTPOLSHPILGMSIOVANDQFYOS  
 FOSEVYLPKEFVTLQPLVCSNMSKLNGLITAKYTKYGRKVDVTLTFLPLSMFK  
 KNTIKTKINGSANFSNDEEMQVNDSSGSLSEYDLSPGVEITITTYTESVGI  
 SRNASTNFFKQHYTILEFPDYTVLKRSLFVIAVTVKTRADGQTLIEERNVAVIT  
 VTQNRVTEYWGSSNGQKMAVOKININTVQSGTFKLEPILDESSLOKAVFLGS  
 KGSMAVHSLFKSPSKTYIOLKTRDENIKVGSPELVVSGNRKLELSPVVSRLQLVA  
 VGNKSNVSLTPENSMTPKACVIVYILEDGEIISDYLKIPVLFPENKIKLYMSKY  
 KAPSEKVSLEISVTPQPSIVGIVAVDXYIDGVVDAEYRFRMEENEGHVDIHDSLSG  
 MNMSPAVFOCGMLVTLDAVTLQVYIDGVVDAEYRFRMEENEGHVDIHDSLSG  
 SPVYRKPPEFTWLDTMGRTIOEFVTPDSITSWVAGFVISELGLITTPV  
 ELQAPQEPFLINLPYSVIRGEERALETTINYLKDAIEVYIIEKSDKPIIMTSGE  
 INATQOQTLVPSDEGATVLPPIRPHLGEIPITVLAISPLASDAVOMLVAAEGL  
 EKSYSOSILDLTNRLOSTLKTLSFSPPTVGSERVOTIAGDVGSPINGLASL  
 IRMPYSGEONMINPAPNIVYLDVLTKKOITLNLKEKALSFMGQVRELLYOREG  
 SPSAFNVDPSGTSWLSFVLRCLFADPYIDIDONTIHRITTYLKHOKSGEMNDP  
 GRVHSELQGNKSPVTLTAVITVLSLYKRYQPNIDVESIHPLSEFSFGISDNT  
 LALITVALSVGNSPKAKELMLTWRAQEGQMOPVWSSEKIDLSQFRLDIEVLA  
 YALISHPLQFOTSEGIPIRMWLSRQNSISFASOTDTVALKLSERPALMNTERTN  
 IOVITVPGSSSPKPLIDTNRLLLOLAEIAYOPAVVINSANGFPAICOLAVVYN  
 VKAGSSRRRSIONOARFDLDVAVKKNLNDVINDVNCBSRQSGRGMALMEVNL  
 LSGMVSERBASISSETYAKVEYDKLNLIDSNVETQFCVNIYAVNFKISNTODAS  
 VSIQVYEPERQAVRSYNSVSKSSCDLCSDVQGRPCEDASGSHHSVIFLFCFK  
 LLVFMELML"  
 BASE COUNT 1709 a 1173 c 1241 g 1772 t  
 ORIGIN  
 Alignment Scores:  
 Pred. No.: 0 Length: 5895  
 Score: 7326.50 Matches: 1427  
 Percent Similarity: 98.82% Conservative: 1  
 Best Local Similarity: 98.75% Mismatches: 0  
 Query Match: 99.71% Indels: 17  
 DB: 6 Gaps: 1  
 US-10-020-095-4 (1-1428) x AX534938 (1-5895)  
 QY 1 MetGlnGlyProProIleuLeuThrAlaAlaIleuIleuCysValCysThrAlaAlaIleu 20  
 DB 113 ATGCAGGCGCCACCGCTCTTAACCGCCCACTCTCTGCTGTGACAGCGCGGCTG 172  
 QY 21 AlaValAlaProGlyProArgPheIleuValThrAlaProGlyIleIleArgProGlyGly 40  
 DB 173 GCCGTGGCTCCCGGCGCTCGGCTTCTGTGACAGCGCCAGAGATCATAGCGCCGGAAGA 232  
 QY 41 AsnValThrIleGlyValGluIleuIleuGluIleuIleuIleuIleuIleuIleuIleu 60  
 DB 233 AATGTGACTATTTGGGCTGAGCTTCTGACACATGCGCTTCAAGTGAAGTGAAGCG 292



QY 61 GluLeuLeuYsThrAlaSerAsnLeuThrValIleuGluValIleGluValIlePhe 80  
 Db 293 GAGCTGCTCAAGACAGACATCAAACTCTCTCTCTCGAAGACGAAGAGCTCTT 352  
 QY 81 GluIlyGlySerPheIlySerThrLeuThrLeuProSerLeuProlAsnSerAlaAspGlu 100  
 Db 353 GAAAGAGGCTCTTTAAAGACCTTACTCTTCATCTACTACTCTGAAACAGCGAGTAC 412  
 QY 101 IleTyrGluLeuArgValIleThrGlyArgThrGlnAspGluIleLeuPheSerAsnSerThr 120  
 Db 413 ATTATAGACTACGTGTACCGGACGTACCGAGATGAGATTTATCTCTTAATAGTACC 472  
 QY 121 ArgLeuSerPheGluThrIlyValArgIleSerValPheIleGlnThrAspIlyAlaLeuTyr 140  
 Db 473 CGCTTATCATTTGAGACCAAGAAATATCTGTCTTCAATTCAAACAGACGAGCTTATAC 532  
 QY 141 LysProIlyGlnIlyValIlyValPheArgIleValIleThrLeuPheSerAspPheIlyProTyr 160  
 Db 533 AAGCCAAAGCAAGAGTAAAGTTTCGATTTGTACACTCTTCTGATTTTAAACCTTAC 592  
 QY 161 LysThrSerLeuAsnIleLeuIleIlyAspProIlySerAsnLeuIleGlnGlnTyrPhe 180  
 Db 593 AAAACCTCTTAAACATTTCTCATTAAGGACCCCAATCAAAATTTGATTCACAGCTGTTG 652  
 QY 181 SerGlnIlySerAspLeuGlyValIleSerIlyThrPheGlnIlySerSerHisProIle 200  
 Db 653 TCACAAACAAGTGAATCTTGAGTCATTTCCAAAATTTTCAGTATCTTCCATCCATA 712  
 QY 201 LeuGlyAspTyrPheIleGlnValGlnValAsnAspGlnThrIlyTyrGlnIlySerPheGln 220  
 Db 713 CTGGTGACTGGTCTATTCAGTTCAGTGAATGATGACAGACATATTCATATCAATCTTTCAG 772  
 QY 221 ValSerGluTyrValIleProIlyPheGluValIleThrLeuGlnThrProlLeuTyrCysSer 240  
 Db 773 GTTTCAGAAATGTATTAACCAAAATTTGAAGTACTTTGACAGACCAATTAATGTTCT 832  
 QY 241 MetAsnSerIlyHisIleAsnGlyIleThrIleThrAlaIlyTyrIleGlyIlyProVal 260  
 Db 833 ATGAAATCTAAGCATTTAAATGTATGATCAACGCAAGTATCAATATGGGAACCGAGT 892  
 QY 261 LysGlyAspValIleThrLeuThrPheLeuProlLeuSerPheTyrGlyIlyValIlyValAsnIle 280  
 Db 893 AAAGAGACGTTACGCTTACATTTTACCTTTATCTTTTGGGAAAGAAAGAAATTAAT 952  
 QY 281 ThrIlyThrPheIlyHisIleAsnGlySerAlaAsnPheSerPheAsnAspGluGluMetIly 300  
 Db 953 ACAAAACATTTAAAGTAATGATCTGCAAACTTCTTTTAAATGAAGAGATGAAA 1012  
 QY 301 AsnValMetAspSerSerAsnGlyLeuSerGluTyrLeuAspLeuSerSerProGlyPro 320  
 Db 1013 AATGTAAATGATTTCTTCAATAGACTTTCTGAATACCTGGATCTATCTTCCCTGGACCA 1072  
 QY 321 ValGluIleLeuThrThrValIleGluSerValIleThrGlyIleSerArgAsnValIleThr 340  
 Db 1073 GTAAATAATTTTAAACACAGTACAGAAATCAAGTACAGGATTTTCAAAATATGTAAGCACT 1132  
 QY 341 AsnValPhePheIlyValIleAspTyrIleIleGluPhePheAspTyrThrThrValIleu 360  
 Db 1133 AATGTGTCTTCAAGCAACATGATTAATCAATCAATGATTTTGTATTAATCACTGCTTGG 1192  
 QY 361 LysProSerLeuAsnPheThrAlaThrValIlyValIleThrArgAlaAspGlyAsnGlnLeu 380  
 Db 1193 AAGCAATCTCTCAACTTCAACGCACTGTGAAGTAACTGCTGTGATGGCAACCAACTG 1252  
 QY 381 ThrLeuGluIlyValArgAsnAsnValIleThrValIleThrGlnArgAsnTyrThrGlu 400  
 Db 1253 ACTTTTAAAGAAAGAAATATGATGATCAATACAGTACACAGAGAAACTATATACGAG 1312  
 QY 401 TyrThrSerGlySerAsnSerGlyAsnGlnIlyMetGluAlaValIleGlnIlyIleAsnTyr 420  
 Db 1313 TACTGAGCGGATCTTAACAGTGAATTCAGAAATGGAAGCTGTTCGAAATTAATTAAT 1372  
 QY 421 ThrValProGlnIlySerGlyThrPheIlyIleGluPheProIleLeuGluIlyAspSerSerGlu 440

Db 1373 ACTGTCCTCCCAAGTGAACCTTTAAGATGAATTCCTCAATCTCGAGGATTCACGTAG 1432  
 QY 441 LeuGlnLeuIlyAlaTyrPheLeuGlySerIlySerSerMetHisAlaHisSerLeuPhe 460  
 Db 1433 CTACAGTGAAGGCTTAATTTCTTGGTAGTAAAGTACGTGACGTCAATGCTGTTT 1492  
 QY 461 LysSerProSerIlyThrTyrIleGlnLeuIlyThrArgAspGluAsnIleIlyValIleGly 480  
 Db 1493 AAGTCTCTTAATGAACATATCATCTCAACTTAAACAGAGATGAATAATTAAGTGGGA 1552  
 QY 481 SerProPheGluLeuValIleSerGlyAsnIlyValSerLeuIlyGluLeuSerTyrMetVal 500  
 Db 1553 TCCCTCTTGAAGTGGTGGTGTAGTGGCAACAAAGATTTGAAGAGCTTATATGTA 1612  
 QY 501 ValSerArgGluGlnLeuValIleValIleGlyIlyValAsnSerThrPheSerLeuThr 520  
 Db 1613 GATTCAGGGGAGAGTGGTGGCTGTAGGAAACAAATTCAAACAAATGTTCTTTAACA 1672  
 QY 521 ProGluAsnSerTyrPheProIlyValIleValIleValIleTyrIleGluAspAspGly 540  
 Db 1673 CAGAAATTTCTTGACTCCAAAGCTGTGTATTTGTATTAATGAAATGATGGG 1732  
 QY 541 GluIleIleSerAspValLeuIlyIleProValGlnLeuValPheIlyAsnIlyIleIly 560  
 Db 1733 GAATTAATAAGTATGTTCTTAAATAATCTGCTGCTCAGCTGTTTAAATAAGATTAAG 1792  
 QY 561 LeuTyrThrSerIlyValIlyValIleGluProSerGluIlyValSerLeuArgIleSerVal 580  
 Db 1793 CTATATTTAGTAAAGTAAAGTGAACCAATCGAAGAAAGTCTCTTGTAGATCTCTGTG 1852  
 QY 581 ThrGlnProAspSerIleValIleGlyIleValAlaValIleAspIlySerValIleLeuMetAsn 600  
 Db 1853 ACAAGCTGATCTTCAATAGTGGATTTAGCTGTTGACAAAGTGTGATCGATGAAT 1912  
 QY 601 AlaSerAsnAspIleThrMetGluAsnValIleHisGluLeuGluLeuTyrAsnThrGly 620  
 Db 1913 GCCTTAATGATTTTCAATGAAATGTGTCATGAGTGGAACTTTAATAACACAGCA 1972  
 QY 621 TyrTyrLeuGluIlyMetPheMetAsnSerPheAlaValIlePheGluGluIlyLeuTyrPval 640  
 Db 1973 TATTAATTTAGGCAATGTTCAATGAATTTCTTTGAGTCTTTCAGAAATGTGACTGTGGTA 2032  
 QY 641 LeuThrAspAlaAsnLeuThrIlyAspTyrIleAspGlyValIlyAspAsnAlaGluTyr 660  
 Db 2033 TTACAGATGCAAACTCAAGAAAGATTAATGATGCTTTATATGACATGACGAATAT 2092  
 QY 661 AlaGluArgPheMetGluGluAsnGluIlyHisIleValAspIleHisAspPheSerLeu 680  
 Db 2093 GCTGAGAGTATTATGAGGAAATGAAGACATATGATGATTCATGACTTTTCTTGG 2152  
 QY 681 GlySerSerProHisValArgIlyHisIlePheProGluThrTyrIleThrLeuAspThrAsn 700  
 Db 2153 GGTGAGAGTCCACATGTCGAAAGCAATTTCCAGACATGGAATTTGGCTGAGCACCAAC 2212  
 QY 701 MetGlyTyrArgIleTyrGlnIlyPheGluValIleThrValProAspSerIleThrSerTyr 720  
 Db 2213 ATGGATTACAGATTTTAAACAAATTTGAAGTACTGTACCTGATCTTACTCTTGG 2272  
 QY 721 ValAlaThrGlyPheValIleSerGluAspLeuGlyLeuGluIleuThrThrThrProVal 740  
 Db 2273 GTGGCTACAGGTTTGTGATCTGAGGAACTGGGCTTGGACTTAACAATACTCCAGTG 2332  
 QY 741 GluLeuGlnAlaPheGlnProPhePheIlePheLeuAsnLeuProTyrSerValIleArg 760  
 Db 2333 GAGCTCAAGCCCTTCAACCAATTTTCATTTTTCATTTTCAATCTTCACTCTGTTATCAGA 2392  
 QY 761 GluGluGluPheAlaLeuGluIleThrIlePheAsnTyrLeuIlyAspAlaThrGluVal 780  
 Db 2393 GGTGAAGATTTGCTTGGAAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2452  
 QY 781 LysValIleIleGluIlySerAspIlyPheAspIleLeuMetThrSerSerGluIleAsn 800

Db 2453 AACGTATCATTCAGAAAAGTGAACAATTGATATTTCTAATGACTTCACAAATGAAATTAAT 2512  
 Qy 801 AATthGlyHISgInglInthLeuLeuValProSerGluAspGlyAlaThrValLeuPhe 820  
 Db 2513 GCCACAGGCCACAGCAGACCTTCTGGTTCCAGAGGAGATGGGCAACTGTTCTTTT 2572  
 Qy 821 ProILeArgProThrHISleuGlyGluLeuProIleThrValIleuLeuSerProThr 840  
 Db 2573 CCCATAGGCCAACACATCTGGAGAAATTCCTATCAAGTCAACACTCTTTCACCCACT 2632  
 Qy 841 AlaSerAspAlaValThrGlnMetIleLeuValIleGluGlyIleGluIlySerSerTy 860  
 Db 2633 GCTTCGATGCTCTCACCCAGATGATTTTGTAAAGGCTGAAAGAAATGAAATAATCAT 2692  
 Qy 861 SerGlnSerIleLeuLeuAspLeuThrAspAsnArgLeuGlnSerThrLeuSerThrLeu 880  
 Db 2693 TCACAAATCATCTTATTAAGCTTGACGTGACAAATAGGCTACAGGTTCCCGAAACCTTGG 2752  
 Qy 881 SerPheSerPheProAsnThrValThrGlySerGluArgValGlnIleThrAlaIle 900  
 Db 2753 AGTTTCTCATTTCTCTCTAATAACAGTCACTGCAAGAAAGATTCAAGATCACTGCAATT 2812  
 Qy 901 GlyAspValIleuGlyProSerIleAsnGlyLeuAlaSerLeuIleArgMetProTyArgI 920  
 Db 2813 GAGATGTTCTTGCTCTTCATCAATAGCTTAAGCTTACCTCATGTGATGGATGCTTAATGC 2872  
 Qy 921 CysGlyGluGlnAsnMetIleAsnPheAlaProAsnIleTyIleLeuAspTyLeuThr 940  
 Db 2873 TGTGGTGAACAGAACATGANTAAATTTTGCTCCAAATTTTACATTTTGATTTCTGACT 2932  
 Qy 941 LysIleLysGlnLeuThrAspAsnLeuLysGluIleAlaLeuSerPheMetArgGlnGly 960  
 Db 2933 AAAAAAACAACCTGACAGATTAATTGAAAGAAAGCTCTTCATTTATGAGCAAGGT 2992  
 Qy 961 TyrGlnArgGlnLeuLeuTyArgIleArgIleAspGlySerPheSerAlaPheGlyAsnTy 980  
 Db 2993 TACCAAGAGAACTTCTCTATCAGAGGAGAGATGGCTTTTCAGTCTTTTGGGAATTTAT 3052  
 Qy 981 AspProSerGlySerThrTrpLeuSerAlaPheValLeuArgCysPheLeuGlnAlaAsp 1000  
 Db 3053 GACCTTCTGGAGACCTTGTTGTCAAGCTTTTGTTAAGATGTTTCTTGAAAGCCGAT 3112  
 Qy 1001 ProTyIleAspIleAspGlnAsnValIleuHISArgThrTyThrTrpLeuLysGlyHIS 1020  
 Db 3113 CCTACATAGATTTATGATCAGATGATGTTACACAGAACATACACTTGGCTTAAGACAT 3172  
 Qy 1021 GlnLysSerAsnGlyGluPheTrpAspProGlyArgValIleHISerGluLeuGlnGly 1040  
 Db 3173 CAGAAATCCAAACGGTAAATTTTGGATCCAGAGAGATGATTCATGATGCTTCAAGGT 3232  
 Qy 1041 GlyAsnLysSerProValThrLeuThrAlaTyIleValThrSerLeuGlnGlyTyArg 1060  
 Db 3233 GGCATTAAGATCCAGTAACTTACAGCTTATATGTAATCTTCTCCGGGAATTAAGA 3292  
 Qy 1061 LysTyGlnProAsnIleAspValGlnGlySerIleHISpHeuGlnSerGluPheSer 1080  
 Db 3293 AAGTACAGCCTTAACATGATGTCAGAGAGTCAATCATTTTGGAGTCTGAATTCAGT 3352  
 Qy 1081 ArgGlyIleSerAspAsnTyThrLeuAlaLeuIleThrTyAlaLeuSerSerValGly 1100  
 Db 3353 AAGAGGAAATTCAGACATTAATCTTACGCCCTTAATCAATTCATCATCAGAGGGG 3412  
 Qy 1101 SerProLysAlaLysGlnAlaLeuAsnMetLeuThrTrpArgIleGlnGlnGlyGly 1120  
 Db 3413 AGTCTTAAGCCGAGAGAGCTTGAATATGCTGACTTGGAGACAAACAGAAAGTGGC 3472  
 Qy 1121 MetGlnPheTrpValSerSerGlySerIleLeuSerAspSerTrpGlnProArgSerLeu 1140  
 Db 3473 ATGCATTTCTGGGTGCATCAGAGTCCAAACTTTCGACCTCCGCGACGACCACTCCCTG 3532  
 Qy 1141 AspIleGluValAlaAlaTyAlaLeuLeuSerHISpHeuGlnPheGlnThrSerGlu 1160  
 Db 3533 GATATTGAAGTTGACACCTATGACATGCTCTCACTTCTTAATTTCAAGCTTCTGAG 3592

Qy 1161 GlyIleProIleMetArgTrpLeuSerArgIleArgAsnSerLeuGlyPheAlaSer 1180  
 Db 3593 GGATCCCAATTATGAGGTGGCTTAAGCAGCAAAAGAAATAGCTGGGTGTTTGATCT 3652  
 Qy 1181 ThrGlnAspThrThrValAlaLeuLysValAlaLeuSerGluPheAlaAlaLeuMetAsnThr 1200  
 Db 3653 ACTCAGGATACCACTGTGGCTTTAAAGGCTGTGTGTAATTTGACAGCCCTTAATGATACA 3712  
 Qy 1201 GluArgThrAsnIleGlnValThrValThrGlyProSerSerProSerPro----- 1217  
 Db 3713 GAAAGCAAAATATCCAAATGACGTTGACGGGCTTACCTCAACAGTCTGTAAGTTT 3772  
 Qy 1218 -----LeuAlaValGlnPro 1223  
 Db 3773 CTGATTGACACACACACCGCTTACTCTTCAGACAGCAGACCTTGCTGTGACGCCA 3832  
 Qy 1224 MetAlaValAsnIleSerAlaAsnGlyPheGlyPheAlaIleCysGlnLeuAsnValVal 1243  
 Db 3833 ATGCAAGTTAAATATTTCCCAATATGTTTGGATTTGCTAATTTGTCAGCTCAATGTTGTA 3892  
 Qy 1244 TyrAsnValIleAlaSerGlySerSerArgArgArgArgSerIleGlnAsnGlnGlnAla 1263  
 Db 3893 TATTAATGTAAGGCTTCTGGGCTTCTTGAAGACCAAGATTCATCCAAATCAAGAGCC 3952  
 Qy 1264 PheAspLeuAspValAlaValIleGlnLeuLysAspAspLeuAsnHISValAspLeuAsn 1283  
 Db 3953 TTTGATTTAGATCTTGCTGTAAAGAAATTAAGATGATCTCAATCATGTGATTTGAAT 4012  
 Qy 1284 ValCysThrSerPheSerGlyProGlyArgSerGlyMetAlaLeuMetGlnValAsnLeu 1303  
 Db 4013 GTGTGTAACAAGCTTTTCCGGCCCGGATGAGATGTCATGCTCTTAATGAAATTAACCTTA 4072  
 Qy 1304 LeuSerGlyPheMetValProSerGluAlaIleSerLeuSerGluThrValIlyLysVal 1323  
 Db 4073 TTAAGTGCTTTATGCTGCTTCAGAGCAATTTCTCGAGGAGAACATGAAAGAGTG 4132  
 Qy 1324 GluTyAspHISGlyLysLeuAsnLeuTyIleLeuAspSerValAsnGlnThrGlnPheCys 1343  
 Db 4133 GAATATGATCAAGAAACTCAACTCTATTATGATTCGTGTAATGAAACCAAGTTTGT 4192  
 Qy 1344 ValAsnIleProAlaValArgAsnPhelyValSerAsnThrGlnAspAlaSerValSer 1363  
 Db 4193 GTTAATATCTCGCTGTGAGAACTTAAAGTTTCAATATCAACAGATGCTTCAAGTTC 4252  
 Qy 1364 IleValAspTyTyGlnProArgArgGlnAlaValArgSerTyAsnSerGluValIly 1383  
 Db 4253 ATAGTGATTAATAAGACCAAGAGACAGCGGCTGAGAAAGTTTAACTCTGAAGTGAAG 4312  
 Qy 1384 LeuSerSerCysAspLeuCysSerAspValGlnGlyCysArgProCysGluAspGlyAla 1403  
 Db 4313 CTGTCTCTCTGACCTTTTGCACTGATGTCAGAGGCTGCCGTCCTTGTAAGATGAGCT 4372  
 Qy 1404 SerGlySerHISHISerSerValIlePheIlePheCysPheLysLeuLeuTyPhe 1423  
 Db 4373 TCAGGCTCCATCAACACACTTTCAGTCATTTTATTTCTGTTCAGAGCTTGTGACTTT 4432  
 Qy 1424 MetGluLeuTrpLeu 1428  
 Db 4433 ATGGAACCTTGGCTG 4447  
 RESULT 5  
 AX537612 4335 bp DNA linear PAT 23-NOV-2002  
 LOCUS AX537612  
 DEFINITION Sequence 3 from Patent WO02070738.  
 ACCESSION AX537612  
 VERSION AX537612.1 GI:25269535  
 KEYWORDS  
 SOURCE  
 ORGANISM Homo sapiens (human)  
 Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.  
 REFERENCE 1

AUTHORS Schuh, A. and Ouwehand, W.  
TITLE Diagnosis and treatment of blood disorders  
JOURNAL Patient : MO 02070738-A 3 12-SEP-2002;  
Schuh, Andre (CA)

FEATURES Location/Qualifiers  
source 1. .4335

/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
1. .>4335  
/note="unamed protein product"

CDS

/protein\_id="CAD57497.1"  
/db\_xref="GI:25269536"  
/translation="MGSPPLTAAHLCCVTAALAVARGPRLVTAGIRPGNVIT  
GVELLHCPQGVTKALIKTASNLTVSVLEAGVPEKSGSKTLTSLPSNADLEY  
ELRTYGRQDEILFNSNTRLSPEKRSISVFTQTKALKKQEKREIVTLFSPKRY  
KTSILNLIKDKPSNLIQOWLSQSDLVISKTQLSHPHIGDWSIQVQVNDQTYOS  
FQSEYVLPRKEVTLQTPLYCSNMSKHLNGITIAKYGKVKQDVLTLPSPFMK  
KKNTTKFKINGSANFSENDENKMDSSNGLESLDLSPPGVEILITVTESVTG  
SRNVTNVEFKQHDYIIEFFDYTLKPSLNFATVKTADGQQLLEERANNVIT  
VTORNTVEYSGNSGNGKMEANOKIKVTVPOSGTPKIEPILIEDSELOKAYFGS  
KSSNAHSLFKSPKTYITQKTRDENIKVSPFELVYSGNKLKELSYMYVSGQLYA  
VGQNSTMPSLTPENSWTPKACVIVYIIEDEGLISVLKIPQVLVEKNKIKLYMSY  
KAPESEKVSLEISVTPDSIVGIVAVKSVNLNANSDITMENVHELEJNTGYLG  
MFMSFAVCEGQAMVLDANLTKQYIDGYVDAVEARFMEENEGHIVDIDSLGS  
SPHYKHPETWILDTNMSRIYQEPFVTPDSTSWATGPFVISEDGLQGLTTTP  
ELQAPQPFETFLNPSVIRGEFALETITFNLYKADTEYVYIIEKDKDILMTSE  
INATGQOTLVPSDEGATVLPFRPHLGIPIVTTALSPADATOTMLVAEET  
EKYSQSIDLDITNRLQSTLKITLSFSEPTVGSERVOITADIVLGSINGLASL  
IRMPYGGCEQMNIAPNIVYLDYLTKKQYLDNLKELKALSPMGQVOREBLVOREB  
SFAFGNDPDSGMSFAFVLCLEADPYIDIDONLHRTYTLKQHSNGSEFMP  
GRVHSELQGNKSPVTLVAYITSLIGRYOPNDIVQSIHFLSESPKGIENDYT  
LAULTLSSYSGSKAKAENMLTMRABOEGMOPWYSSSKLSDSPSPSIDSEVA  
YALISHPLOFTSGEIPIMRWLSRORNSLGEFASDTDTVALKALSPFALMDERTN  
IQVTVTGSSSPVKFLIDTNRLQLQTAELVQPMANISANGFPAICQLVAVN  
VKAGSSRRRSIONOEAFLDIDVAVKENDLHNDLVNCTSPGPERSGALMEVNL  
LSGPMVSEASLSEYKVEYDHGKMLYLDSENETOFCVNI PAVENFKVSNODS  
VSIYDYEPBRQAVRSYNSVKKSSCOLCSVQGCRCPDGASGSHHSVIFPCKK  
LTYMEML."

BASE COUNT 1295 a 838 c 922 g 1280 t  
ORIGIN

## Alignment Scores:

Pred. No.: 0 Length: 4335  
Score: 7317.50 Matches: 1426  
Percent Similarity: 98.75% Conservative: 1  
Best Local Similarity: 98.69% Mismatches: 17  
Query Match: 99.58% Indels: 1  
DB: 6 Gaps: 1

US-10-020-095-4 (1-1428) x AK537612 (1-4335)

QY 1 MetGlnGlyProProLeuLeuThrAlaAlaHisLeuLeuCysValCysThrAlaAlaLeu 20  
Db 1 ATGAGAGGCGCCACCGCTCTGACCGCGCCACCTCTCTCGTGTGACGCGCGCTG 60  
QY 21 AlaValAlaProGlyProAlaGlyPheLeuValThrAlaProGlyIleIleArgProGly 40  
Db 61 GCCGTGGCTCCCGGCGCTCGGTTCTGTGTGACGCGCCAGGAGCATCAGCGCCGAGGA 120  
QY 41 AsnValThrIleGlyValGluLeuLeuGluHisCysProSerGlnValIleThrValAla 60  
Db 121 AATGTGACTATTGGGAGTGGAGCTTCTGTGAACCTGCGCTTACAGGTGACTGTAAAGCG 180  
QY 61 GluLeuLeuYsrThrAlaSerAsnLeuThrValSerValLeuGluAlaGluGlyValPhe 80  
Db 181 GAGTGTCTCAAGACAGCATCAACCTCTCTGTCTGTCTGTGAAGACGAAGAGTCTTT 240  
QY 81 GluYsrGlySerPheLeuYsrThrLeuThrLeuProSerLeuProLeuAsnSerAlaAspGlu 100  
Db 241 GAAAGAGGCTCTTTAAGACACTTACTCTTCATCACTACTGAAACAGTGCAGATGAG 300

QY 101 IleYsrGluLeuArgValThrGlyArgThrGlnAspGluIleLeuPheSerAsnSerThr 120  
Db 301 ATTATGAGCTACGTGTAAACCGGACGATCCAGAGATGAGATTTATCTCTAATAGTACC 360  
QY 121 ArgLeuSerPheGluThrIleYsrGlyIleSerValPheIleGlnThrAspGlyAlaLeuTyr 140  
Db 361 CGCTTACATTTAAGACCAAGAAATATGCTGTTCATTCATAACAGCAAGGCGCTTATAC 420  
QY 141 LysProLeuGlnGluValLysPheArgIleValIleThrLeuPheSerAspPheLysProTyr 160  
Db 421 AAGCCAAAGCAAGATGAGAGTTCCGATCTGTACACTCTTCCAGATTTAAGCCTTAC 480  
QY 161 LysThrSerLeuAsnIleLeuIleLysAspProLysSerAsnLeuIleGlnIleTyrLeu 180  
Db 481 AAAACCTCTTAAACATTCCTAATTAAGACCCCAATTCAAATTTGATCCAAAGTGGTTG 540  
QY 181 SerGlnGlnSerAspLeuGlyValIleSerLysThrPheGlnLeuSerSerHisProIle 200  
Db 541 TCACAAACAAAGTATCTTGAGTCATTTCCAAAACCTTTCAGCTATCTCCATCCAAATA 600  
QY 201 LeuGlyAspTrpSerIleGlnValAlaAsnAspGlnThrTyrTyrGlnSerPheGln 220  
Db 601 CTGCTGACTGCTCTATTCAAGTTCAGATGAATGACCAAGACATATATCATCTTTACG 660  
QY 221 ValSerGluTyrValLeuProLysPheGluValThrLeuGlnThrProLeuTyrCysSer 240  
Db 661 GTTTCGAATATGATTAATCAAAATTTGAAGTCACTTGTGACACATATATATTTGTT 720  
QY 241 MetAsnSerLysHisLeuAsnGlyThrIleThrAlaLysTyrThrTyrGlyLysProVal 260  
Db 721 ATGAATTTCAAGCATTTAAATGATGATCAATCAAGCAAGATTCATATGGAAGCCAGTG 780  
QY 261 LysGlyAspValThrLeuThrLeuProLeuSerPheTrpGlyLysLysAsnIle 280  
Db 781 AAGGAGACCTTAAGCTTATTAATTTTACCTTTATCTTTGGGAAAGAGAAATATT 840  
QY 281 ThrLysThrPheLysIleAsnGlySerAlaAsnPheSerPheAsnAspGluIleTyr 300  
Db 841 ACAAAAACATTTAAGATTAATGATGATCGCAAACTTCTCTTATATGATGAAGATGAAA 900  
QY 301 AsnValMetAspSerSerAsnGlyLeuSerGluTyrLeuAspLeuSerProGlyPro 320  
Db 901 AATGTATGATTCCTTCAATATGATGATTCGATATCTGATATCTTCCCTGACCA 960  
QY 321 ValGluIleLeuThrThrValThrGlnSerValThrGlyIleSerArgAsnValSerThr 340  
Db 961 GTTGAATTTTAAACACAGACAGACAGATCAGTTACAGGTATTTCAAGAAATGATGACAT 1020  
QY 341 AsnValPhePheLysGlnHisAspTyrIleIleGluPheAspTyrThrThrValLeu 360  
Db 1021 AATGTGTTCTTCAAGCAACATGATTTACATCATTAAGTTTTTATTAATTAATCTGCTTG 1080  
QY 361 LysProSerLeuAsnPheThrAlaThrValLysValThrArgAlaAspGlyAsnGlnLeu 380  
Db 1081 AAGCCATCTCTCAATTCACAGCAGCAGTGAAGTAACTGTCCTAATGCAACCAACTG 1140  
QY 381 ThrLeuGlnGluArgArgAsnAsnValIleIleThrValThrGlnArgAsnTyrThrGlu 400  
Db 1141 ACTCTTGAAGAAAGAAATTAATGTATGATCACTGACACAGAGAACTTATCTGAG 1200  
QY 401 TyrTrpSerGlySerAsnSerGlyAsnGlnLysMetGluAlaValGlnLysIleAsnTyr 420  
Db 1201 TACTGAGCGGATCTTAACAGTGAATTCGAATTCGAATTCGAATTCGAATTCGAATTC 1260  
QY 421 ThrValProGlnSerGlyThrPheLysIleGluPheProIleLeuGluAspSerSerGlu 440  
Db 1261 ACTGTCCTCCCAAGTGAAGTCTTTAAGATTTGATTCATCCATCTGCGAGATTCAGAG 1320  
QY 441 LeuGlnLeuValAlaTyrPheLeuGlySerLysSerSerMetAlaValHisSerLeuPhe 460  
Db 1321 CTACAGTTGAAGGCTTATTTCTGTGTGATGATGAAGTGCAGTTCATCTGTTT 1380  
QY 461 LysSerProSerLysThrTyrIleGlnLeuLysThrArgAspGluAsnIleLysValGly 480

Db	1381	AAAGTCTCTAGTAAGACATACATCCAACTTAAAAACAAGAGTGAATAATTAAGGTGGGA	1440
OY	481	SerProPheGluLeuValValSerGlyAsnLysArgLeuLysGluLeuSerTyrMetVal	500
Db	1441	TGCGCTTTTGAGTTGGTGGTTAGTGGCAACAACAGATTGAAGAGATTAAAGCTATAGGTA	1500
OY	501	ValSerArgGlyGlnLeuValAlaValGlyLysGlnAsnSerThrMetPheSerLeuThr	520
Db	1501	GTATCCAGGGGACAGTTGGTGGCTGTAGGAAAAAATAATTCACACATGTTCTCTTAAACA	1560
OY	521	ProGluAsnSerTyrThrProValAlaCysValIleValTyrTyrIleGluAsnAspGly	540
Db	1561	CCAGAAATTTCTTGACATCCCAAAACCTGTGTAAATTTGGTATTATTAAGAAATGATGGG	1620
OY	541	GluIleIleSerAspValLeuLysIleProValGlnLeuValPheLysAsnLysIleLys	560
Db	1621	GAAATTATTAAGATGATGTTCTTAAAAATTCGTGTACGTTGTTTTTAAAAATTAAGTAAAG	1680
OY	561	LeuTyrTyrPheSerLysValLysAlaGluProSerGlyLysValSerLeuArgIleSerVal	580
Db	1681	CTATATTGTGAGTAAGTAAGTAAAGCTGAACCATCTGAGAAAGTCTCTTAAAGCATCTCTGG	1740
OY	581	ThrGlnProAspSerIleValGlyIleValAlaValAlaLysSerValAsnLeuMetAsn	600
Db	1741	ACACAGCTCATCTCATAGTGGAGTTGAGTGTGTGACAAAGTGTAAATCTGATGAT	1800
OY	601	AlaSerAsnAspIleThrMetGluAsnValValHisGluLeuGluLeuTyrAsnThrGly	620
Db	1801	GCCTTATATGATATTAACAATGGAATAATGTGGTCATGATGATGGACCTTTATACACAGGA	1860
OY	621	TyrTyrLeuGlyMetPheMetAsnSerPheAlaValPheGlnLysCysGlyLeuThrVal	640
Db	1861	TATATTATTAGGATGTTGATGATGATCTTTTGGACGCTTTCAGAGATGTGAGACTCTGGGTA	1920
OY	641	LeuThrAspAlaAsnLeuThrLysAspTyrIleAspGlyValTyrAspAsnAlaGluTyr	660
Db	1921	TTGACAGATGCAAACTCAGCAAGGATTATATGTAGTGTGTTAAGACAAAGCAGAAAT	1980
OY	661	AlaGluArgPheMetGluLysAsnGluGlyHisIleValAspIleHisAspPheSerLeu	680
Db	1981	GCTGAGAGCTTTATGAGAGAAAATTAAGACATATTAAGATTAATCAGATCTTTCTTGG	2040
OY	681	GlySerSerProHisValArgLysHisPheProGluThrTyrIleThrLeuAspThrAsn	700
Db	2041	GGTACAGATCCACATGTCGGAAGCATTTTCCAGACCTTGATTTGGCTAGACCAAC	2100
OY	701	MetGlyTyrPheGlyIleTyrGlnGluPheGluValThrValProAspSerIleThrSerTyr	720
Db	2101	ATGGGTTCCAGGATTTACCAAGAAATTTAAGTAACGTAACTGATTCATCACTTCTTGG	2160
OY	721	ValAlaThrGlyPheValIleSerGluAsnLysGlyLeuGlyLeuThrThrThrProVal	740
Db	2161	GTGGCTACTGGTTTGTATCTCTGAGGACCTGGGCTTTGGACATAACAACATCACTCCAGTG	2220
OY	741	GluLeuGlnAlaPheGlnProPhePheIlePheLeuAsnLeuProTyrSerValIleArg	760
Db	2221	GAGCTCCAGGCTTCCAAACATTTTCAATTTTTTTTGAATCTTCCCTACTCTGTATACGA	2280
OY	761	GlyGluGluPheAlaLeuGluIleThrIlePheAsnTyrLeuLysAspAlaThrGluVal	780
Db	2281	GGTGAAGAATTTGCTTTGGAAATACTATATTCATTTTGAAGAAGTCCACGTAGGTT	2340
OY	781	LysValIleIleGluLysSerAspLysPheAspIleLeuMetThrSerSerGluIleAsn	800
Db	2341	AAAGTAATCATTAAGAAAAAGTACAAATTTGATTTCTTAAGACTTCAAAAGAAATAAAT	2400
OY	801	AlaThrGlyHisGlnGlnThrLeuLeuValProSerGluAspGlyAlaThrValLeuPhe	820
Db	2401	GGCAAGGCGCACACACACCTTCTGGTCCCATGAGATGGGGCAACTGTTCTTTTT	2460
OY	821	ProIleArgProThrHisLeuGlyGluIleProIleThrValThrAlaLeuSerProThr	840

Db	2461	CCCATCAGGCCAACACACACTGGGAGAAATTCCTATCACGTCACAGCTCTTTCACCCACT	2520
QY	841	AlaserApalavalThrgImecileuVallysalaglugilyileuglyuSerYr	860
Db	2521	GCTTCGTAGCTGTACCCACAGATATTTTAAAGGCTGAAGAAATGAAATAATCATAT	2580
QY	861	SerGInserIleleuLeuAspLeuThraSpahrgleugInserThrleuYsthrleu	880
Db	2581	TCACAAATCCAACTTAATAGACTTCACGCAATAGGCTACAGAGATACCTGTAACATTTG	2640
QY	881	SerPheSerPheProProAsnThrValThrcGlySerGluArgValGlnIleThAlalle	900
Db	2641	AGTTTCATTTCTCTCTTAATACGTACTGGCAGTGAAGAAGTTACAGATCACTGCATTT	2700
QY	901	GlyAspValleugIlyProSerIleasnGlyLeuAlaserLeuIleargMetProIyGly	920
Db	2701	GGAGATGTTCTTGGCTTCATCAATGGCTTACCTTCATTTGATTCGGATCCCTTATGGC	2760
QY	921	CysglYglugInasMetIleasnPheAlaProAsnIleYrIleleuAspYrleuThr	940
Db	2761	TGTGTGTAAACGAACATGATTAATTTTGCTCCAAATATTTTACATTTTGGATTACTGACT	2820
QY	941	LysIlyslYsgIlnleuThraSpasrleuYslYslYsalaleuSerPheMetArgGlnIly	960
Db	2821	AAAAAGAAACACACTGCACAGATTAATTTGAAAGAAAGCTCTTCAATTTATAGGCAAGGT	2880
QY	961	TyrgInargGluLeuLeuYrGlnArgGluAspGlySerPheSerAlaPheGlyAsnYr	980
Db	2881	TACCAAGAGAACTTCTCTATCAGAAGGAATGGCTCTTCATGCTCTTTGGGAATATT	2940
QY	981	AspProSerGlySerThrTrpLeuSerAlaPheValleuArgCysPheLeuGluAlasp	1000
Db	2941	GACCTTCCTGGAGACACTGGTGTGTACGCTTTTGTTTAAGATGTTTCTTGAAGCCGAT	3000
QY	1001	ProYrIleAspIleAspGlnAsnValleuIleIargThrYrThrTrpLeuYglYHis	1020
Db	3001	CCTTACATAGATATTGATCAGAATGTGTACACGAACATACCTTGGCTTTAAAGAAAT	3060
QY	1021	GlnYysSerAngIyGluPheTrpAspProGlyArgValIleHisSerGluLeuGlnIly	1040
Db	3061	CAGAAATCCACGGTGAATTTTGGGATCCAGGAAGAGTATTCATAGAGCTTCAGGT	3120
QY	1041	GlyAsnYysSerProValThrLeuThraIlyrIleValThrSerLeuLeuGlyYrArg	1060
Db	3121	GGCAATAAAGATCCAGTAACACTTACAGCCTATATTGTAATCTTCCTCTGGATATATGA	3180
QY	1061	LysYrGlnProAsnIleAspValGlnGluSerIleHisPheLeuGluSerGluPheSer	1080
Db	3181	AAGATATCGCTTAACATTTAGTGTCCAGAGCTCTATCCATTTTGGAGTCTGAATTCAGT	3240
QY	1081	ArgGlyIleSerAspAsnYrThrLeuAlaleuIleThrYrAlaleuSerSerValGly	1100
Db	3241	AGAGGAATTTAGCAAAATTAATCTAGCCTTAAATTAATTAATGCAATGTCAATCAGTGGG	3300
QY	1101	SerProYsalalYsgIlnaleuAsnMetLeuThrTrpArgAlagIugInglYglY	1120
Db	3301	AGTCCTTAAGCGAAGGAAGCTTTGAATATGTGACTTGGAGACGACGAACAAGAGGTGGC	3360
QY	1121	MetGlnPheTrpValSerSerGluSerIlyslYleuSerAspSerTrpGlnProArgSerIeu	1140
Db	3361	ATGCAAATTTCTGGGTGTATCAGAGTCCAAATCTTCTGACTCTCGGACGCCAGCTCCCTG	3420
QY	1141	AspIleGluValAlaAlaYrAlaleuLeuSerHisPheLeuGlnPheGlnThSerGlu	1160
Db	3421	GATATGTAGATTGACGCTATGACCTGTCTCACACTTTCACATTTTCAGACTTCTGAG	3480
QY	1161	GlyIleProIleMetArgTrpLeuSerArgGlnArgAsnSerLeuGlyIlyPheAlaser	1180
Db	3481	GGATTCCAAATTAATGAGGTGTGCTACGACGCAAGAAATAGCTTGGGTGGTITTGCACT	3540
QY	1181	ThrgInAspThrThrValAlaleuYsalaleuSerGluPheAlaAlaleuMetAsnThr	1200
Db	3541	ACTCAGGATACACTGTGTGGCTTTAAAGCTGTGTCTAAATTTGACGCCCTATATGAATACA	3600

QY 1201 GUAAGTThraenileglnValThrValThnGlyProSerSerProSerPro----- 1217  
 Db 3601 GAAAGACAAATATCCAGTAGTACCCTGACGGGGCTGAGCTCACCAGAGCTCTTAAGTTT 3660  
 QY 1218 -----LeuAlaValAlaGlnPro 1223  
 Db 3661 CTGATTGACACACAAACCGCTTACTCTCTTACAGACAGAGAGCTTGGCTGCTGCTGACGCCA 3720  
 QY 1224 MetAlaValAlaenileSerAlaasnGlyPheGlyPheAlaIleCysGlnLeuAsnValVal 1243  
 Db 3721 ATGGCAGTTAATATTTCCGCAAAATGTTTGGATTGCTATTGCTACCTCAATGTTGTA 3780  
 QY 1244 TyrAsnValIlyAlaSerGlySerSerArgArgArgSerIleGlnAsnGlnAla 1263  
 Db 3781 TATATGGAAGGCTTGGGCTCTTCAAGAGACAGAGATCTATCCAAATCAAGAAAGCC 3840  
 QY 1264 PheAsnLeuAsnValAlaValIlyGlnAlaValAsnValAsnValAsnValAsn 1283  
 Db 3841 TTTGATTAGATGTTGCTGTAAAGAAATAAAGATGATCTCAATCAATGATGATTTGAAT 3900  
 QY 1284 ValCysThrSerPheSerGlyProGlyArgSerGlyMetAlaLeuMetGlnValAlaAsn 1303  
 Db 3901 GTGCTTACACCTTTTCCGGCCCGGGTACGAGTGGCTCTTATGGAAGTTAACCTTA 3960  
 QY 1304 LeuSerGlyPheMetValProSerGlnAlaIleSerLeuSerGlyThrValIlyAlaVal 1323  
 Db 3961 TTAAGTGGCTTTAAGTGCTCTTCAAGAGCAATTTCTGACGAGACAGCTGAAGAAAGTG 4020  
 QY 1324 GlnTyrAsnIlyGlyValLeuAsnLeuTyrLeuAsnSerValAlaGlnThrGlnPheCys 1343  
 Db 4021 GAATATGATCATGGAAGAACTCAACCTCTATTCAATCTGTAAATGAAGAACCTTTTGT 4080  
 QY 1344 ValAsnIleProAlaValAlaArgAsnPheValSerAsnThrGlnAsnAlaSerValSer 1363  
 Db 4081 GTTAATATTTCTGCTGTGAGAACTTTAAAGTTTCAATTACCAGAGATGCTTCACTGCTTC 4140  
 QY 1364 IleValAsnTyrTyrGlnProArgArgGlnAlaValAlaArgSerTyrAsnSerGlnValIly 1383  
 Db 4141 ATATGAGTATCTATGAGCCAGAGAGACAGCGGTGAGAACTTCAACTCGAAGTGAAG 4200  
 QY 1384 LeuSerSerCysAsnLeuCysSerAsnValAlaGlnIlyCysArgProCysGlnAsnGlyAla 1403  
 Db 4201 CTGTCCCTCTGTGACCTTTGACAGTATGTCAGAGGCTGCTCTTGTGAGAGATGAGAGCT 4260  
 QY 1404 SerGlySerIleAsnIleSerSerValIlePheIlePheCysPheValLeuTyrPhe 1423  
 Db 4261 TCAGGCTCCATCATCACTCTTCACTCTTTTATTTTCTGTTCAAGCTTCTTACTTT 4320  
 QY 1424 MetGlnLeuTyrLeu 1428  
 Db 4321 ATGGAACCTTGGCTG 4335  
 RESULT 6  
 AX534936 4761 bp DNA linear PAT 22-NOV-2002  
 LOCUS Sequence 3 from Patent WO02070696.  
 DEFINITION AX534936  
 ACCESSION AX534936  
 VERSION AX534936.1 GI:25261478  
 KEYWORDS  
 SOURCE  
 ORGANISM Homo sapiens (human)  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 REFERENCE 1  
 AUTHORS Schuh, A. and Sutherland, R. D.  
 TITLE Cdl09 nucleic acid molecules polypeptides and methods of use  
 JOURNAL Patent: WO 02070696-A 3 12-SEP-2002;  
 Schuh, Andre (CA); Sutherland, Robert D. (CA)  
 FEATURES  
 source 1..4761  
 location/Qualifiers  
 /organism="Homo sapiens"  
 /mol\_type="genomic DNA"

CDS  
 /db\_xref="taxon:9606"  
 113..4450  
 /note="unnamed protein product"  
 /codon\_start=1  
 /protein\_id="CAD57232.1"  
 /db\_xref="GI:25261479"  
 /translation="MQGPPPLTLPAHLVCYCTAALAVAPGPRELVTAPEGIRREGNAVIT  
 GVELLHPCSOQVYVKAELKLTASNLVSLAEAGVEKGSFKTLTLPSPJLNSADBIY  
 ELAVRTODEILFNSNTRLSFETKISYFIOTDKALYRKQEFKRIYTLFEDFKPY  
 KSLNLTLDKPSNLTQWLSQSDGVSKTFOLASHPLDGMWSIQOVNDQDTYOS  
 FOYSEYVLPKFEVTLTQPLVCGSNKHLNCTTAKTQKPKVGDVTLPLSPFGK  
 KKITTKTFKINSANFSFNDEKKNVMSNGSLSELDSSPEPVILITVTSVYIT  
 SRNVSTNVPFKQHDYILFFDYTVLAKPPLNLFVTRADENQTLTERRNNVIT  
 VTQRNTEVSGNSNGNQMEAVOKLNYVPOGTEKIFPPLIEDSELQKAYFLGS  
 KSMVAHSLFKSPSKTYIOLKTRDENIKYGSFELVSGNKLKELSYWVSGQLVA  
 VGKONSTWLSLPENSTPRACVYIYVIEDDGEISDYIKIPOLVFKKIKLYMKRG  
 KAPRSKVLRLSVTOPDSITVGIVADKSVNLMNSDITMENVVELLYNNGYIG  
 MPNAPVAVQECCLAVLTANLTQYIDYVDAEAEKPFMEBECHYIDIDHPSLGS  
 SPVRRGHPEPWTWLDITNNGSRIYQEFVTVPDISTSWATGVLISEDGLGTTTPV  
 ELQAPQPEIPLNLPYSVIRGEFALERTIPYLNKDAVEVKYIEKSDKFDILMTNE  
 INATGHOQLLVSEEDGATVLPPIRTHLGEIPIYVTLASPASDVTOMIILVKEGI  
 EKSYSOILLDLTNDRLSTLTKLSPSPNTVTGSEKVOITAGVLPISNGLASL  
 IRMPVCGERONMINFAPNTYIIDLTKKQILDNLKERKLSMRQGOBELLYOREDG  
 SRSAFENYPSGSLTSLAFVLCFPLEADYIIDQVNLRTYTWLGHKRSNGEFPD  
 GRAVHSELQGNKSPVTLVAYLVTSLGFRKYOPNIDVGSJHFLSEFSRGSIDNYT  
 LALITVALSVGSPKAKELNMLTWABEGGQFVWSSESKLSDSQPRSLDIIEVA  
 YALISHLFOFSEGIPIRMRLSRQNSLGEFASITDDTVALKALSEFALAMTERTN  
 IOVTVGPSPSPVKFLIDTHNRLLIOTLAEVAVOMAVNISANGKSEFICOLNVVYN  
 VASGSSRRRRRSIQNEAFDIDVAYKENDDLNHDVLANVCTFSGSGRSGMLMEVNL  
 LSGFVPSSEALISSEVTKVEYDHGKLNLYLDSVETQCVNIPAVNRKVSNTQDAS  
 VSLVDVPRQAVRSYNSVSKSLCDLSDVQGCRCPCDGASGSHHSVIFIFPCFK  
 LLYFMELMTL"  
 BASE COUNT 1419 a 913 c 1016 g 1413 t  
 ORIGIN  
 ALIGNMENT Scores:  
 Pred. No.: 0 Length: 4761  
 Score: 7317.50 Matches: 1426  
 Percent Similarity: 98.75% Conservative: 1  
 Best Local Similarity: 98.69% Mismatches: 17  
 Query Match: 99.58% Indels: 17  
 DB: 6 Gaps: 1  
 US-10-020-095-4 (1-1428) x AX534936 (1-4761)  
 QY 1 MetGlnGlyProProleuLeuThrAlaIleLeuLeuCysValCysThrAlaIleLeu 20  
 Db 113 ATCGAGGGCCACACGCTCTGACCGCGCCACCTCTCTGCTGTGACCGCGGCTG 172  
 QY 21 AlaValAlaProGlyProArgPheLeuValThrAlaProGlyIleIleArgProGlyIly 40  
 Db 173 GCCGTGGCTCCGGGCTCGGTTCTGTGTGACAGCCCGAGGATCATGAGCGCGGAGGA 232  
 QY 41 AsnValThrIleGlyValGlnLeuLeuGlnIlyAsCysProSerGlnValThrValIlyAla 60  
 Db 233 AATGTACTATTGGGTGAGCTTCTGGAACCTGCTCCCTTCAACAGCTGATGGAAGGGC 292  
 QY 61 GlnLeuLeuIlyThrAlaSerAsnLeuThrValSerValLeuGlnAlaGlnIlyAlaPhe 80  
 Db 293 GAGCTGCTCAAGACACATCAACCTCACTGCTCTGTCTGTGAGACAGAGAGAGCTTT 352  
 QY 81 GlnIlyGlySerPheIlyThrLeuThrLeuProSerLeuProleuAsnSerAlaAsnGln 100  
 Db 353 GAAAGAGCTCTTTTAAAGACATTACTTCTTCACTATCACTCTTGAACAGTGAGATGAG 412  
 QY 101 IleTyrGlnLeuArgValThrGlyArgThrGlnAsnGlnIleLeuPheSerAsnSerThr 120  
 Db 413 ATTATGACCTACGTATACCGGACGTACCGAGATGAATTTATTTCTTATATGATAC 472  
 QY 121 ArgLeuSerPheGlnThrIlyArgIleSerValPheIleGlnThrAspIlyAlaLeuTyr 140  
 Db 473 CGCTTATCAITTTGACACCAAGAAATATCTGCTTCAATTCACACAGACAGGCTTATAC 532



QY 141 LysProlysglnGluValLysPheAspGlyLeuAlaThrLeuPheSerAspPheLysProTyr 160  
 Db 533 AACCCAAAGCAAAAGGAAAGTTTCGATGTTACCTCTTCAGATTTTAAAGCCCTTAC 592  
 QY 161 LysThrSerLeuSnnlleuLleLysAspProLysSerAsnLeuIleGlnGlnTrrPhe 180  
 Db 593 AAAACCTCTTAAACATTTCAATTAAGACCCCAAAATTCAAATTTGATCCAAAGTGATG 652  
 QY 181 SerGlnGlnSerAspLeuGlyValIleSerLysThrPheGlnLeuSerSerHisProIle 200  
 Db 653 TCACAAACAAGGTATCTTGAGATTCATTTCCAAAACCTTTCAAGTATCTTCCCATCCAA 712  
 QY 201 LeuGlyAspTrrSerIleGlnValGlnValAsnAspGlnThrTyrTyrGlnSerPheGln 220  
 Db 713 CTGGGAGACGTGCTATTTCAAGTTCAAGTAAGACACAGCATATATCATCATTTCCAG 772  
 QY 221 ValSerGlnTrrValLeuPProLysPheGlnValThrLeuGlnThrProLeuTyrCysSer 240  
 Db 773 GITTCAGAATATATATTATACCAAAATTTGAAGTACCTTGACACCATTAATATGTTCT 832  
 QY 241 MetAsnSerLysHisLeuSnnlleLysThrIleThrAlaLysTyrThrTyrGlyLysProVal 260  
 Db 833 ATGAATTCCTAGACATTTAAATGATACATACCGCAAAAGTATCATATGGAAAGCCAGTG 892  
 QY 261 LysGlyAspValThrLeuThrPheLeuProLeuSerPheTrrGlyLysLysAsnIle 280  
 Db 893 AAAGGAGACGTACGCTTACATTTTATCTTATCCCTTTGGGAAAGAAATAAT 952  
 QY 281 ThrLysThrPheLysIleAsnGlySerAlaAsnPheSerPheAsnAspGlnGluMetLys 300  
 Db 953 ACAAAACATTTAAGATAAGATGATCGCAACCTTCTTTAATATGAAGAGATGA 1012  
 QY 301 AsnValMetAspSerSerAsnGlyLeuSerGlnTyrLeuAspLeuSerSerProGlyPro 320  
 Db 1013 AATGTAATGATTCCTTCAATGATGATCTTGGAATCTGATCTTATCTTCCCTGACCA 1072  
 QY 321 ValGlnIleLeuThrThrValThrGlnSerValThrGlyIleSerArgAsnValSerThr 340  
 Db 1073 GTAGAATTTTAAACACAGTACAGATCAGATCAGATTTCAAGATTTCAAGAAATGTAACACT 1132  
 QY 341 AsnValPhePheLysGlnHisAspTyrIleIleGlnPhePheAspTyrThrValLeu 360  
 Db 1133 AAGTGTCTTCAAGAACATGATTCATCATTTGAGTGTGATTAATTAATTAATTAATTA 1192  
 QY 361 LysProSerLeuSnnlleLysThrAlaThrValLysValThrArgAlaAspGlyAsnGlnLeu 380  
 Db 1193 AACCACTCTCACTTCAAGCACAGTACAGAGTAACTGTCGTAATGGCAACCACTG 1252  
 QY 381 ThrLeuGlnGlnArgArgAsnAsnValIleThrValThrGlnArgAsnTyrThrGln 400  
 Db 1253 ACTCTTGAAGAAAGAAATATATGATGATACATACAGTACAGAGAACTTACTGAG 1312  
 QY 401 TyrTrrSerGlySerAsnSerGlyAsnGlnLysMetGlnAlaValGlnLysIleAsnTyr 420  
 Db 1313 TACTGAGGCGATCTAACAGTGAATAATCAGAAATATGAGTGTTCAGAAAATAATTAAT 1372  
 QY 421 ThrValProGlnSerGlyThrPheLysIleGlnPheProIleLeuGlnLysSerSerGln 440  
 Db 1373 ACTGTCCCAAGAGTAACTTTTAAGATTGAATTCCAATCTGAGAGATTCACATGAG 1432  
 QY 441 LeuGlnLeuLysAlaTyrPheLeuGlySerLysSerSerMetAlaValHisSerLeuPhe 460  
 Db 1433 CTACAGTTGAAGCCATATTTCTTGATGTAAGAAAGTACATGACATTCATAGTCTGTT 1492  
 QY 461 LysSerProSerLysThrTyrIleGlnLeuLysThrArgAspGlnAsnIleLysValGly 480  
 Db 1493 AAGTCTCCATGAAGCATATCAATCCAACTTAAACAAAGATGATAAATATTAAGCGGGA 1552  
 QY 481 SerProPheGlnLeuValValSerGlyAsnLysArgLeuLysGlnLeuSerTyrMetVal 500  
 Db 1553 TCCTCTTTGAGTTGGTGGTGGTGGCAACAAACGATGAAGAGTTAAGCTATATAGTA 1612  
 QY 501 ValSerArgGlyGlnLeuValAlaValGlyLysGlnAsnSerThrMetPheSerLeuThr 520

Db 1613 GTATCCAGGGGACAGTGGTGGCTGAGAGAAACAAATTCACAAATGTTCTTTAACAA 1672  
 QY 521 ProGlnAsnSerTrrThrProLysAlaCysValIleValTyrTyrIleGlnLysAspGly 540  
 Db 1673 CCAAGAAATTTCTGGATCTCCAAAGCCTGTGTAATGTTGATTAATGAATGAATGAGG 1732  
 QY 541 GluIleIleSerAspValLeuLysIleProValGlnLeuValPheLysAsnLysIleLys 560  
 Db 1733 GAAATTTATAGATGATCTTAAATAATCTCTGTACCTGTTTAAAAATTAAGATTAAG 1792  
 QY 561 LeuTyrTrrSerLysValLysAlaGlnProSerGlnLysValSerLeuArgIleSerVal 580  
 Db 1793 CTATATGGAGTAAGTGAAGTGAACCTGACATCTGAGAAAGTCTCTTGAATCTCTGG 1852  
 QY 581 ThrGlnProAspSerIleValGlyIleValAlaValAspLysSerValAsnLeuMetAsn 600  
 Db 1853 ACACAGCCTGATCCATAGTGGATGTAATGTAATGTAATGTAATGTAATGTAAT 1912  
 QY 601 AlaSerAsnAspIleThrMetGlnAsnValValHisGlnLeuGlnLeuTyrAsnThrGly 620  
 Db 1913 GCTCTATGATATTAATGAAATGTAATGTAATGTAATGTAATGTAATGTAATGTAAT 1972  
 QY 621 TyrTyrLeuGlyMetPheMetAsnSerPheAlaValPheGlnLysCysGlyLeuTrrVal 640  
 Db 1973 TATTAATTTAGCATGTTGATGAATTTCTTTGCACTTTACAGAAATGTAATGTAAT 2032  
 QY 641 LeuThrAspAlaAsnLeuThrLysAspTyrIleAspGlyValTyrAspAsnAlaGluTyr 660  
 Db 2033 TTGACAGATGCAAACTCAAGAAAGATTAATATATGATGTTATGACAAATGCAAAAT 2092  
 QY 661 AlaGlnArgPheMetGlnGlnLysGlnGlyHisIleValAspIleHisAspPheSerLeu 680  
 Db 2093 GCTGAGAGTTTATGAGAGAAATGAAGACATATGTAATGTAATGTAATGTAATGTAAT 2152  
 QY 681 GlySerSerProHisValArgLysHisPheProGlnTrrTrrIleTrrPheAspThrAsn 700  
 Db 2153 GGTACAGTCCACATGTCGAAAGCATTTTCCAGAGACTTGATTTGGTACACACCAAC 2212  
 QY 701 MetGlyTyrArgIleTyrGlnGlnPheGlnValThrValProAspSerIleThrSerTrr 720  
 Db 2213 ATGGGTCCAGGATTTTACCAAGAAATTTGAAGTAATGTAATGTAATGTAATGTAAT 2272  
 QY 721 ValAlaThrGlyPheValIleSerGlnAspLeuGlyLysGlyLeuThrThrPheProVal 740  
 Db 2273 GTGGCTACTGGTTTGTGATCTGAGACCTGGGCTTGGACTGACATCACTCCAGTG 2332  
 QY 741 GlnLeuGlnAlaPheGlnProPhePheIlePheLeuAsnLeuProTyrSerValIleArg 760  
 Db 2333 GAGCTCCAAAGCTTCCAAACATTTTCAATTTTGTGAATCTTCCCTACTCTGTATCAGA 2392  
 QY 761 GlyGlnGlnPheAlaLeuGlnIleThrIlePheAsnTyrLeuLysAspAlaThrGlnVal 780  
 Db 2393 GGTGAAGAAATTTGCTTGGAAATACTATATTTCAATTTTGAAGATGCCACTGAGGT 2452  
 QY 781 LysValIleIleGlnLysSerAspLysPheAspIleLeuMetThrSerSerGlnLysAsn 800  
 Db 2453 AAGTAATCATTTGAAGAAAGTGAATAATTTGATATTTCAATGACTTCAATGAATAAT 2512  
 QY 801 AlaThrGlyHisGlnGlnThrLeuLeuValProSerGlnLysAspGlyAlaThrValLeuPhe 820  
 Db 2513 GCCACAGGCCACACACACACCTTCTGTTGCCATGAGATGGGGCAACTGTTCTTTT 2572  
 QY 821 ProIleArgProThrHisLeuGlyGlnIleProIleThrValThrAlaLeuSerProThr 840  
 Db 2573 CCCATCAGGCCCAACATCTGGAGAAATTTCTATCAACAGTCAAGCTCTTCAACCACT 2632  
 QY 841 AlaSerAspAlaValThrGlnMetIleLeuValLysAlaGlnGlyIleGlnLysSerTyr 860  
 Db 2633 GCTTCTGATGCTGACCAAGATGATTTTGAAGGCTGAGAGATGAAGAAATCAATAT 2692  
 QY 861 SerGlnSerIleLeuLeuAspLeuThrAspAsnArgLeuGlnSerThrLeuLysThrLeu 880

Db 2693 TCACAAATCCATCTTAATTAAGCTTGAACGATAGGCTACAGAGTACCCTGAAAACTTTG 2752  
 Qy SerPheSerPheProAsnThrValThrGlySerGluArgValGlnIleThrAlaIle 900  
 Db 2753 AGTTTCATCTTCTCTCAATACAGTGACTGGCAGTGAAGAGTTTCAGATACCTGCAATT 2812  
 Qy GlnAspValIleuGlyProSerIleAsnGlyLeuAlaSerLeuIleArgMetProTyrGly 920  
 Db 2813 GGAATGTTCTTGGTCTTCATCAATGGCTTACCTCATTTGATTCGGATCCCTTATGGC 2872  
 Qy CysGlyGluGlnAsnMetIleAsnPheAlaProAsnIleTyrIleLeuAspTyrLeuThr 940  
 Db 2873 TGTGTGAACAGAACATGATTAATTTGTCTCAATATTTTCAATTTTGGATTAATCTGACT 2932  
 Qy LysIleValSerGlnLeuThrAspAsnLeuGlyLysValAlaLeuSerPheMetArgGlnGly 960  
 Db 2933 AAAAAGAAACAACTGACAGATTAATTTGAAAGAAAAGCTCTTTCAATTAATGAGCAAGGT 2992  
 Qy TyrGlnArgGluLeuLeuTyrGlnArgGluAspGlySerPheSerAlaPheGlyAsnTyr 980  
 Db 2993 TACCAGAGAGAACTTCTCTATCAAGAGGAGATGGCTCTTTCAGTGTCTTTGGCAATTAT 3052  
 Qy AspProSerGlySerThrTyrPheLeuSerAlaPheValLeuArgCysPheLeuGlnAlaAsp 1000  
 Db 3053 GACCCCTCTGGAGACATTTGGTGTCTGAGCTTTTGTAAAGATTTTCTTGAAGCCGAT 3112  
 Qy ProTyrIleAspIleAspGlnAsnValIleuHISArgThrTyrThrTyrPheLeuGlyValHis 1020  
 Db 3113 CTTTACATAGATATTGATCAGAAATGTGTTACAGAAACATACCTTGGCTTAAAGCAAT 3172  
 Qy GlnLysSerAsnGlyGluPheTyrAspProGlyArgValIleHisSerGluLeuGlnGly 1040  
 Db 3173 CAGAAATCCAAACGGGATTTTGGGATCCAGAGAGATGATTAATGAGCTTCAAGGT 3232  
 Qy GlnAsnLysSerProValThrLeuThrAlaTyrIleValThrSerLeuLeuGlyTyrArg 1060  
 Db 3233 GGCATATAAAGCTCAAGTAACCTTAACGCTTATTTGAACTTCTCCCTGGGATATAAGA 3292  
 Qy LysTyrGlnProAsnIleAspValGlnGluSerIleHisPheLeuGluSerGluPheSer 1080  
 Db 3293 AAGTATCAAGCTTAACATTTGATGTGCAAGATCTATTCATTTTGGAGTCTGAATTCAGT 3352  
 Qy ArgGlyIleSerAspAsnTyrThrLeuAlaLeuIleThrTyrAlaLeuSerSerValGly 1100  
 Db 3353 AGAGAAATTTCAAGCAATTAATACCTGACCTTAATTAATGATGATGATGATGATGATG 3412  
 Qy SerProLysAlaLysGluAlaLeuAsnMetLeuThrTyrPheAlaGlnGlnGlyGly 1120  
 Db 3413 AGTCTTAAGCGAAGGAAAGCTTTGAATATGCTGACTTGGAGACAGAAAGAAAGTGGC 3472  
 Qy MetGlnPheTyrValSerSerGluSerTyrLysSerAspSerTyrGlnProArgSerLeu 1140  
 Db 3473 ATGCAATTTCTGGGTGTCACTACAGATCCAAACTTTCTGACTCTGACAGCCAGCTCCCTG 3532  
 Qy AspIleGluValAlaAlaTyrAlaLeuLeuSerHisPheLeuGlnPheGlnThrSerGlu 1160  
 Db 3533 GATATTAAGAGTTGACGCTTAAGCACTGCTCTCAACATTTCTTAACATTTCAAGATTTCTGAG 3592  
 Qy GlnTyrIleProIleMetArgTyrPheLeuSerArgGlnArgAsnSerLeuGlyGlyPheAlaSer 1180  
 Db 3593 GGAATCCCAATTAATGAGGTGGCTTAAGCAGCGAAAGAAATAGCTTGGGTGGTTTGGCATCT 3652  
 Qy ThrGlnAspThrThrValAlaLeuLysAlaLeuSerGluPheAlaAlaLeuMetAsnThr 1200  
 Db 3653 ACTAGAGATACACTGTGGCTTAAAGCTCTGTCTGAATTTGACAGCCCTTAATGAATACA 3712  
 Qy GlnArgThrAsnIleGlnValThrValThrGlyProSerSerProSerPro----- 1217  
 Db 3713 GAAAGCAAAATATCAAGATACCGTGAACGGGGCTTACCTCAACAACTCCGTAAAGTTT 3772  
 Qy 1218 -----LeuAlaValGlnPro 1223  
 Db 3773 CTGATGACACACAAACCGCTTACTCTTCAAGACAGAGCTTGTCTGTGTACAGCCA 3832

Qy 1224 MetAlaValAsnIleSerAlaAsnGlyPheGlyPheAlaIleCysGlnLeuAsnValVal 1243  
 Db 3833 ATGCAAGTTATATTTCCGAAATGGTTTGGATTTTGCATTTTGTACAGCTCAATGTTGTA 3892  
 Qy 1244 TyrAsnValIleValAsnGlySerSerArgArgArgSerIleGlnAsnGlnGlnAla 1263  
 Db 3893 TATTAATGTGAAGCTTCTGGGTCTTCTCAAGAGCGAAGATCTTATCCAAATCAAGAGCC 3952  
 Qy PheAspLeuAspValAlaValLysGluAsnLysAspAspLeuAsnHisValAspLeuAsn 1283  
 Db 3953 TTTGATTTACATTTTGTCTGTAAGAAATTAAGATGATCTCAATCATGTGGATTTGAAT 4012  
 Qy ValCysThrSerPheSerGlyProGlyArgSerGlyMetAlaLeuMetGlnValAsnLeu 1303  
 Db 4013 GTGTGTAACAAGCTTTTGGGCCCGGGTAGAGATGGATGGCTTTATGGAATTAACCTTA 4072  
 Qy LeuSerGlyPheMetValProSerGluAlaIleSerLeuSerGluThrValLysValVal 1323  
 Db 4073 TTAAATGGCTTTTATGGTGGCTTCAAGAGCAATTTCTCTGACGACAGATGAAGAAAGTG 4132  
 Qy GlnTyrAspHisGlyLysLeuAsnLeuTyrLeuAspSerValAsnGlnThrGlnPheCys 1343  
 Db 4133 GAATATGATCATGAAACTCAACCTTATTGATTTCTGTAATGAAACCCAGTTTGT 4192  
 Qy ValAsnIleProAlaValArgAsnPheLysValSerAsnThrGlnAspAlaSerValSer 1363  
 Db 4193 GTTATATTTCTCTCTGAGAACTTTAAAGTTTCAATTAATCCAAATGCTTCAGTGTCC 4252  
 Qy IleValAspTyrTyrGluProArgArgGlnAlaValArgSerTyrAsnSerGluValLys 1383  
 Db 4253 ATAGTGATTAATCATAGAGCCAAAGAGACAGCGGTGAGAAATTAACAATCTTGAAGCAAG 4312  
 Qy LeuSerSerCysAspLeuCysSerAspValGlnGlyCysArgProCysGluAspGlyAla 1403  
 Db 1384 TTTTCTCTCTGATCCTTGTGAGTGAATGTCAGAGGCTGCGCTTGTGAGATGAGAGCT 4372  
 Qy 4313 CTGTCTCTCTGATCCTTGTGAGTGAATGTCAGAGGCTGCGCTTGTGAGATGAGAGCT 4372  
 Db 1404 SerGlySerHisHisSerSerValIlePheIlePheCysPheLysLeuTyrPhe 1423  
 Qy 4373 TCAGGCTCCCATATCATCACTTTCATTTTATTTTCTGTTTCAAGCTTCTGTACTTT 4432  
 Qy 1424 MetGlnLeuThrPhe 1428  
 Db 4433 ATGGAATTTGGCTG 4447  
 Qy Db 4433 ATGGAATTTGGCTG 4447  
 RESULT 7  
 AX534940 LOCUS 5895 bp DNA linear PAT 22-NOV-2002  
 AX534940 DEFINITION Sequence 7 from Patent WO02070696.  
 AX534940 ACCESSION  
 AX534940.1 GI:25261486 VERSION  
 KEYWORDS  
 SOURCE  
 ORGANISM  
 Homo sapiens (human)  
 Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE  
 1 Schuh, A. and Sutherland, R.D.  
 Cationic nucleic acid molecules polypeptides and methods of use  
 Patent: WO 02070696-A 7 12-SEP-2002;  
 Schuh, Andre (CA) ; Sutherland, Robert D. (CA)  
 Location/Qualifiers  
 1..5895  
 /organism="Homo sapiens"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9606"  
 113..4450  
 /note="unnamed protein product"  
 /codon\_start=1  
 /protein\_id="CAD57234.1"  
 /db\_xref="GI:25261487"  
 /translation="MGGPPLTPAHLVCYCTAALAVAPGRFLVTPAGIIRPGNVTI  
 GVELLEHCPQVTVKALKLTASNLTVSLAEAGVFEKSPFTLTLPISPLNSADBIY



ELRVTRQDDEILFNSNTRLSFETKRIISVFIQTDKALYRKQKQEVKRIIVTLFSDEKPY  
KTSNLIILKDPKSNLIQOWLQQSDGLVSKTFOLSHPLIGWMSIQOVNDQTYOS  
PQVSEYVLPKFEVTLQTPLYCSMNSKHLNGITAKTAKPYKPVGDTVTLFPLSFMGK  
KNKIKTFKINGSANSFNDEEMKQWMSNGLSYLDLSPSPVTLITTVESVGI  
SNVSTNVFVKOHDIIEFPDPTVYLKSLNFTATYKTRPAGNOCLTEBRANVIT  
VYORATYEMSSGNSGNQKMEAVOKINNTVYPOSGTKIERPILDESSEIQAKYRIGS  
KSMVAHSLFKSPSKTYIQLKRDENIKVGSPELVSNKRIKELSYNVSRGQIYA  
VGNKONSTMSLPENSWTPKACVIVYIIDDEIISDVAKIPIQULFKKIKLYMSKV  
KAPSEKVSRLISVTPDSIVGIVADKVNLMANANDIMENVAHELNTYGYLG  
MFNSPFAVFOEGAVLTDLNLTQYIDGYNDABYAEKPMENBEHUIIDHFSIGS  
SPHVRKHPEETWMLDMMNGSRVYQSEFTVDSITSWATGVISEDGLGTTTPV  
ELQAFOPFIPLNLPYSVIRGEFALEITTFNLTQATKATVAKIENSKRFDILMNSI  
INATGHOQTLVPESEDATVLEPIRPHLGEIPIVTYALSPASADVTMMIIVKAGI  
KESYSQSIILDLTDNRLOSTLKSFSPPNVTGSERYQTALIGVLFPSINGLASL  
IRMPYCGEONMINFAPNIYILDYLFKKOLTDNLKELKLSFMROGYORLELYOREJG  
SFEAFGNVDPGSGTWLSAFVLRCELEADYIIDIQNVLTARTYTWLKGOKSGNEFDP  
GRVTHSELOGKNSPYTLNAYITSLGTRKTOPINDVDSIHFLSESREKISDNTY  
LALITVALLSVGSFKAKELNMLTWBAEDGQMFWVSSEKLSDSWQPSLIDIEVA  
YALSHPLQFQTSSEGIPIRMWLSRONSJGGFASJODTVALKALSEFALNTEKTN  
IQVTVGPSSPSPVKFLIDTNRLILQTLAEVAVCPMAVNI SANGGFALCOLNVVN  
VKAAGSSRRRSIIONQEAEDLVAVAKENDLNHVDLNTCTSPSGRSGMALMEVNI  
LSEFWPSRAILSETVKKVEYDHGKLNLYLDSVNETOCVNI PAVRNKSNVTOAS  
VSIYDIYERPRQAVRSYNSSEVKLSCDLSDVQGCPCEDGASGSHHSSVIFICEK  
LITFMEML"

BASE COUNT 1708 a 1174 c 1241 g 1772 t  
ORIGIN

## Alignment Scores:

Pred. No.: 0 Length: 5895  
Score: 7317.50 Matches: 1426  
Percent Similarity: 98.75% Conservative: 1  
Best Local Similarity: 98.69% Mismatches: 1  
Query Match: 99.58% Indels: 17  
DB: 6 Gaps: 1

US-10-020-095-4 (1-1428) x AXS34940 (1-5895)

QY 1 MetGInGlyProProLeuLeuThraAlaAlaHleLeuLeuCySValCySThraAlaAlaLeu 20  
DB 113 ATGCAAGGCGCCACCGCTCTCGACCGCGCCACCTCTCGCTGCTGCGACCGCGCGCTG 172  
QY 21 AlaValAlaProGlyProProPheLeuValThraAlaProGlyTleIleArgProGlyGly 40  
DB 173 GCCGTGCTCCCGCGGCTGCTTCTGCTGACAGCCCGACGATCATCAGCGCCGAGGA 232  
QY 41 AenValThraIleGlyValGluLeuLeuGluHleCySProSerGlnValThraValIleVala 60  
DB 233 AATGTACTAATGGGTGGAGCTTCTGGAACTGCTGCTTCAACAGGTGACTGTGAAGGGG 292  
QY 61 GluLeuLeuYSerThraIleSerAenLeuThraValSerValIleuGluIleGluGlyValaPhe 80  
DB 293 GACCTCTCAAGACAGATCAACCTCACTGCTCTGCTCGAGAGGAGAGAGCTTT 352  
QY 81 GluYLeuGlySerPheYSerThraLeuThraLeuProSerLeuProLeuAenSeraIleAenGlu 100  
DB 353 GAAAAAGGCTCTTTAAAGACCTTACTCTTCACTCACTGTAACAGTGAAGTAG 412  
QY 101 IleTyrGluLeuArgValThraGlyArgThraGlnAspGluIleLeuPheSerAenSerThra 120  
DB 413 ATTATGAGCTACGTGAACCGGACGTACCCAGAGAGAGATTATTTCTCTAATAGTACC 472  
QY 121 ArgLeuSerPheGluThraIleYSerArgIleSerValPheIleGluThraAspIleValaLeuTyr 140  
DB 473 CGCTTATCATTTAGACCAAGAGAAATATCTGCTTCACTCAACCAAGACAGGCTTATAC 532  
QY 141 LysProIleGlnIleValLysPheArgIleValThraLeuPheSerAspPheLysProTyr 160  
DB 533 AAGCCAAAGCAAGAGAGATTTCGATGTTACACTCTTCCAGATTTTAAAGCTTAC 592  
QY 161 LysThraSerLeuAenIleLeuIleLysAspProLysSerAenLeuIleGlnGlnTyrPleu 180  
DB 593 AAAACCTCTTAAACATTCATTAAAGAGACCCCAAAATCAAAATTTGATCCAAAGTGGTGG 652

QY 181 SerGInGInSerAspLeuGluValIleSerIleThraPheGlnLeuSerSerIleProIle 200  
DB 653 TCACAAACAAAGGATCTTGAGTCAATTCCTTCCAAACCTTTTCAAGCTATCTCCATCCAAATA 712  
QY 201 LeuGlyAspTyrSerIleGlnValGlnValAenAspGlnThraTyrTyrGlnSerPheGln 220  
DB 713 CTTGGGATCGGTCTTATTCACATTCACAGGAATGACAGACATATTAATCAATTCATTCAG 772  
QY 221 ValSerGluTyrValIleuProLysPheGluValThraLeuGlnThraProLeuTyrCySser 240  
DB 773 GTTTAGAAATATGATTTACCAAAATTTGAGGACTTTCACACACCATTAATATGTTCT 832  
QY 241 MetAenSerIleYSerGlnIleThraIleThraAlaYSerThraTyrGlyLysProVal 260  
DB 833 AAGAAATTCACATTTAAATGTACCATCAAGGCAAGTATACATATGGAAGCAGAG 892  
QY 261 LysGlyAspValThraLeuThraPheLeuProLeuSerPheTyrGlyLysIleYSerAenIle 280  
DB 893 AAAGAGAGAGTAAACCTTACATTTTACCTTATCTTTTGGGAAAGAAATAATAT 952  
QY 281 ThrIleThraPheIleIleAsnGlySerIleAenPheSerPheAenAspGluIleuLeuIys 300  
DB 953 ACAAAAACATTTAAAGATTAATGATCTGCAAACTTCTTTAAATGATGAAGAGATGA 1012  
QY 301 AenValMetAspSerSerAenGlyLeuSerGluTyrIleuAspLeuSerSerProGlyPro 320  
DB 1013 AATGTAAATGATTTCTTCAATAGACTTCTGTAATACGTGATCTTCCCTGAGCA 1072  
QY 321 ValGluIleLeuThraThraValThraGluSerValThraGlyIleSerArgAenValSerThra 340  
DB 1073 GTAGAAATTTTAAACACAGTGAACAAATCAGTTAAGGATTTCAAGAAATGTAAGCACT 1132  
QY 341 AenValPhePheIleGlnIleAspTyrIleIleGluPhePheAspTyrThraThraValLeu 360  
DB 1133 AATGTGCTTCAACCAACATGATTAATCATCATGATTTTGTGATTAATCATCTGCTTG 1192  
QY 361 LysProSerLeuAspPheThraIleThraValIleYSerValThraArgAlaAspGlyAenGlnLeu 380  
DB 1193 AAGCATCTCTCAATCTTCAACGCCACTGGAAGGTACCTGCTGATGAGCAACCACTG 1252  
QY 381 ThrLeuGluGluArgArgAenAenValIleThraValThraGlnArgAenTyrThraGlu 400  
DB 1253 ACTCTTGAAGAAAGAAATATATGATGATCAATACATGACAGTGAAGAAATATATCTGAG 1312  
QY 401 TyrTyrPserGlySerAenSerGlyAenGlnIleYSerMetGluAlaValGlnIleAenTyr 420  
DB 1313 TACTGAGCGGATCTTAACAGTGAATCAAGAAATGGAAGCTGTTCAAGAAATTAATATAT 1372  
QY 421 ThraIleProGlnSerGlyThraPheIleIleGluPheProIleLeuGluIleuAspSerGlu 440  
DB 1373 ACTGTCCCCCAAGTGAACCTTTTAAGATTGAATTCCTCAATCTTGAGGATTCAGTAG 1432  
QY 441 LeuGlnLeuYSerIleThraPheLeuGlySerIleYSerSerMetAlaValIleSerLeuPhe 460  
DB 1433 CTACAGTTGAAGGCTTATTTCTGTTAGTAAAGATGAGATGCGATCATGTCTGTTT 1492  
QY 461 LysSerProSerIleYSerThraIleGlnLeuYSerThraGlnAspGluAenIleYSerValGly 480  
DB 1493 AAGTCTCTCAATGAACATACATCACTCACTCACTCACTCACTCACTCACTCACTCACT 1552  
QY 481 SerProPheGluLeuValIleSerGlyAenIleYSerGluLeuYSerGluLeuSerIleMetVal 500  
DB 1553 TCGCTTTTGAAGT 1612  
QY 501 ValIleSerArgIleGlnLeuValAlaValGlyIleYSerGlnAenSerThraPheSerLeuThra 520  
DB 1613 GTATCAAGGAGACAGT 1672  
QY 521 ProGluAenSerTyrThraProLysAlaCySValIleValIleTyrTyrIleGluAspAspGly 540  
DB 1673 CCGAAGAAATTTCTGAGCTCCAAAGCCTGTGTAAATGTGTAAATGTGAATGATATGGG 1732  
QY 541 GluIleIleSerAspValLeuYSerIleProValGlnLeuValPheYSerAenIleYSer 560

```

Db      1733 GAAATTAATAGTGGTCTTAAATAATTCCTGTCAGCTGTTTAAATAATAATAAG 1792
Qy      561 LeuTYrTPSerLysValLysAlaGluProSerGluLysValSerLeuArgLLeuSerVal 580
Db      1793 CTATATTTGAGCTTAAGTGAAGCTGACCACTGAGAAAGCTCTCTTGGATCTCTG 1852
Qy      581 ThrGlnProAspSerLLeuAlGlyLLeuAlaValAspLysSerValAsnLeuMetAsn 600
Db      1853 ACACAGCCTGACCTCATGTGTTGGGATTTAGCTGTGCAAAAGTGAATCTGATGAAT 1912
Qy      601 AlaSerAsnAspLLeuThMetGluAsnValValHISGluLeuGluLeuTYrAsnThGly 620
Db      1913 GCCCTTAATGATATTACATGAAATGTAATGTGTTCATGATGGAACCTTTAATACAGGA 1972
Qy      621 TYrTYrLeuGlyMetPheMetAsnSerPheAlaValPheGlnGluCysGlyLeuTYrVal 640
Db      1973 TATATTTTGGCAGTCTTCAATGATTTCTTTGACAGCTTTCCAGAAATGTGGACTGGGTA 2032
Qy      641 LeuThrAspAlaAsnLeuThrLysAspTYrLLeuAspGlyValTYrAspAsnAlaGluTYr 660
Db      2033 TTGACAGATGCAAACTCAGCAAGGATTAATGATGTTTATGACAAATGCAAGAAATAT 2092
Qy      661 AlaGluArgPheMetGluGluAsnGluGlyHISLLeuAlaAspLLeuAspPheSerLeu 680
Db      2093 GCTAGAGGTTTATGAGGAAATGAAAGACATATGTAATATTCATGACTTTTCTTTG 2152
Qy      681 GlySerSerProHISValArgLysHISpHeProGluTYrLLeuTYrLLeuAspThAsn 700
Db      2153 GGTTCAGCTCCAGTCCAGAACGATTTTCCAGAGCTTGGATTTGGCTTACAGCAAC 2212
Qy      701 MetGlyTYrArgLLeuTYrGlnGluPheGluValThValProAspSerLLeuThSerTYr 720
Db      2213 ATGGGTTCCAGATTTTACCAAGAAATTTGAACTGAACTGATTCATCATCTTTG 2272
Qy      721 ValAlaThrGlyPheValLLeuSerGluAspLeuGlyLeuGlyLeuThThProVal 740
Db      2273 GTGGCTACTGTTTGTGATCTCTGAGAGCCTGGGCTTGGACTTCAACACTCCAGTG 2332
Qy      741 GluLeuGlnAlaPheGlnProPhePheLLeuAsnLeuProTYrSerValLLeuArg 760
Db      2333 GAGCTCCAGCTCCAGCAACCTTTTCTATTTTGTGATCTTCCCTACTCTGTTATCGA 2392
Qy      761 GlyGluGluPheAlaLeuGluLLeuThLLeuPheAsnTYrLeuLysAspAlaThGluVal 780
Db      2393 GGTGAAGAATTTGCTTGGAAATACTATTCATTAATTGAAAGATGCCAGAGGTT 2452
Qy      781 LysValLLeuLLeuGlyLysSerAspLysPheAspLLeuMetThSerSerGluLLeuAsn 800
Db      2453 AAGGTATTCATTTGAGAAAGTGAACAATTTGATATTCATTAATGACTTCAAAATGAATAAAT 2512
Qy      801 AlaThrGlyHISGlnGlnThrLeuLeuValProSerGluAspGlyValAlaThValLeuPhe 820
Db      2513 GCCACAGCCACAGCAACCTTCTGCTTCCAGTGGAGTGGGCACTGTTCTTTT 2572
Qy      821 ProLLeuArgProThHISLeuGlyLLeuLLeuProLLeuThValThAlaLeuSerProTh 840
Db      2573 CCATTCAGGCCAACAACATCTGGAGAAATTCCTATTCACAGTCAAGCTTTTCAACCACT 2632
Qy      841 AlaSerAspAlaValThrGlnMetLLeuValLysAlaGluGlyLLeuLysSerTYr 860
Db      2633 GCTTCTGATGCTGCACCAAGATGATTTTAAAGGCTGAAGAAATGAATAATCATAT 2692
Qy      861 SerGlnSerLLeuLeuAspLeuThrAspAsnArgLeuGlnSerThrLeuLysThLeu 880
Db      2693 TCACAAATTCATTTATTAAGCTTGACTGACAAATAGGCTACAGATACCTGAAAACTTTG 2752
Qy      881 SerPheSerPheProAsnThValThGlySerGluArgValGlnLLeuThAlaLLeu 900
Db      2753 AGTTTCATTTCTCTCTTAATACAGTACGCGAGTGAAGATTCAGATACCTGCAATT 2812
Qy      901 GlyAspValLLeuGlyProSerLLeuAsnGlyLeuAlaSerLeuLLeuArgMetProTYrGly 920

```

```

Db      2813 GAGATGTTCTTGCTCCCTCCATCAATGGCTTAGCTCATTTGATTCGATCCCTTATGGC 2872
Qy      921 CysGlyGluGlnAsnMetLLeuAsnPheAlaProAsnLLeuTYrLLeuAspTYrLeuTh 940
Db      2873 TGTGGTGAACAGAACTGATTAATTTTGTCTCCAAATATTTTACATTTTGGATTATCGACT 2932
Qy      941 LysLysLysGlnLeuThrAspAsnLeuLysGluLysAlaLeuSerPheMetArgGlnGly 960
Db      2933 AAAAAGAAACACATGACAGATTAATTTGAAAGAAAAGCTCTTTCATTTATGAGCAAGGT 2992
Qy      961 TYrGlnArgGluLeuLeuTYrGlnArgGluAspGlySerPheSerLLeuPheLysAsnTYr 980
Db      2993 TACCAGAGAACTCTCTATTCAGAGGAGATGGCTTTTCACTGCTTTTGGAAATTAAT 3052
Qy      981 AspProSerGlySerThrTYrLeuSerAlaPheValLeuArgCysPheLeuGlnAlaAsp 1000
Db      3053 GACCTTTCTGGAGCACTTGTTGTGAGCTTTGTTTAAGATGTTTCCTTGAAGCGAT 3112
Qy      1001 ProTYrLLeuAspLLeuAspGlnAsnValLLeuHISArgThTYrThrTYrLeuLysGlyHIS 1020
Db      3113 CTTTACATGATATTTGATCAGAAATGTGTTACACAGAACTATCACCTGGCTTAAAGACAT 3172
Qy      1021 GlnLysSerAsnGlyGluPheThAspProGlyArgValLLeuHISerGluLeuGlnGly 1040
Db      3173 CAGAAATCCAAACGTAATTTTGGATCCAGAGAGTGAATTAATGTAAGCTTCAAGGT 3232
Qy      1041 GlyAsnLysSerProValThrLeuThAlaTYrLLeuAlaThSerLeuLeuGlyTYrArg 1060
Db      3233 GGCATTAATAAGTCCAGTAACACTTACAGCTTAATATTTGTAATCTTCTCCGGGATTAAGA 3292
Qy      1061 LysTYrGlnProAsnLLeuAspValGlnGluSerLLeuHISpHeLeuGluSerGluPheSer 1080
Db      3293 AAGTATCAGGCTTAACATTTGATGCAAGAGTCAATCATTTTGGAGTCTGAATTCAGT 3352
Qy      1081 ArgGlyLLeuSerAspAsnTYrThrLeuAlaLeuLLeuThTYrAlaLeuSerSerValGly 1100
Db      3353 AGAGAAATTCACACATTAATCTTACGCCCTTAATACTTATGATATGTCATCATGAGGGG 3412
Qy      1101 SerProValAlaLysGlnAlaLeuAsnMetLeuThTYrArgAlaGluGlnGlnGly 1120
Db      3413 AGTCTTAACGCAAGAGAGCTTTGAATATGCTGACTTGGAGAGCAACAAGAGGTGGC 3472
Qy      1121 MetGlnPheThTYrLysSerSerGluSerLysLeuSerAspSerTYrGlnProArgSerLeu 1140
Db      3473 ATCAATTTCTGGGTGTCATCAGAGTCCAAACTTTCTGACCTCCGAGCAGCACCCTCCCTG 3532
Qy      1141 AspLLeuValAlaAlaTYrAlaLeuLeuSerHISpHeLeuGlnPheGlnThrSerGlu 1160
Db      3533 GATATTTGAAGTCCAGCTTATGCACTGCTTCCACACTTCTTACAAATTTCAAGACTTCTGAG 3592
Qy      1161 GlyLLeuProLLeuMetArgTYrLeuSerArgGlnArgAsnSerLeuGlyGlyPheAlaSer 1180
Db      3593 GGAATCCCAATTTATGAGTGGCTTAACAGCAAGCAAGAAATAGCTTGGGTGTTTGCATCT 3652
Qy      1181 ThrGlnAspThrThValAlaLeuLysAlaLeuSerGluPheAlaAlaLeuMetAsnTh 1200
Db      3653 ACTCAGATTCACACTGCTGCTTAAAGGCTCTGCTGAATTTGCAACCCCTAATGAATACA 3712
Qy      1201 GluArgThAsnLLeuGlnValThValThArgLysProSerSerProSerPro----- 1217
Db      3713 GAAAGGACAAATTCACAGTGAACGCGGAGGCTTAAGTCAACAAGCTCTGTAAGATT 3772
Qy      1218 -----LeuAlaValLLeuGlnPro 1223
Db      3773 CTGATGACACACACAAACGCTTACTCTTACAGACAGAGAGCTTCTGTGTGTAACGCCA 3832
Qy      1224 MetAlaValAsnLLeuSerAlaAsnGlyPheGlyPheAlaLLeuCysGlnLeuAsnVal 1243
Db      3833 ATGCGAGTTAATATTTCCGAAATGTTTTGATTTGCTAATTTGACAGCTCAATGTTGTA 3892
Qy      1244 TYrAsnValLysAlaSerGlySerSerArgArgArgSerLLeuLysGlnGlnGlnAla 1263
Db      3893 TATTAATGTAAGGCTTCTGTGGTCTTCTAGAAAGCAAGATCTATCCAAATCAAGAGGCC 3952

```

QY 1264 pheaspleuaspvala1a1lysgluasnllyasaspseuanshisvala1aspseuans 1283  
1284 valcysrthserseuerglyp-proglya-rsgserglymecalaleuwtetgluvala1leu 1303  
1304 leuSerGlypHemEValProserGluVala1leSerleuSerGluThrVallyleVal 1323  
1324 GluTyraSpHisGlylyleuansleuTyrlleuaspserVala1nsgluThrglnpheCys 1343  
1344 valasnlepevala1a1adganphelevala1setranthrglnaspalaservValSer 1363  
1364 llevalaspTyrtYrGluProArgArgGlna1a1ValArgserTyranSerGluVal 1383  
4253 ATAGTGAATTACTTAAGCCAGAGACAGCGCGGTGAAGATTACAACTTGAGAGAG 4312  
1384 leuSerSerCysaspseuCyseraspVala1glnGlyCysArgProCysglua1a1a 1403  
4313 CTGTCTCTCTGACCTTTCGACGTGATGTCACAGGCTGCGCTCTGTGAGATGAGCT 4372  
QY 1404 SerGlySerHisHisSerSerVal11lephelepheCysphelyleuLeuTyrrpe 1423  
4373 TCAGGCTCCCATCATCACTTCTGATCATTTTATTTCTGTTCAAGCTTCTGACTTT 4432  
QY 1424 MetGluLeuThrepu 1428  
4433 ATGGAACCTTGGCTG 4447  
Db

RESULT 8  
AX534942 3535 bp DNA linear PAT 22-NOV-2002  
LOCUS  
DEFINITION  
Sequence 9 from Patent WO02070696.  
AX534942  
VERSION  
AX534942.1 GI:25261490  
KEYWORDS  
SOURCE  
Homo sapiens (human)  
ORGANISM  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE  
AUTHORS  
Schub, A. and Sutherland, R.D.  
JOURNAL  
Cd109 nucleic acid molecules polypeptides and methods of use  
Patent: WO 02070696-A 9 12-SEP-2002;  
Schub, Andre (CA) ; Sutherland, Robert D. (CA)  
FEATURES  
source  
1. .3535  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
113. .3316  
/note="unnamed protein product"  
/codon\_start=1  
/protein\_id="CAD57235.1"  
/translation="MGGPPLTRAHLLCTCTAALAVAPGPRPLVTAIPGIIRGNVTI  
GVELLEHCSQVTVKALKTASNLTVSLAEVGEKSFKTLPLSLMSADELY  
ELAVTGRTODEILFNSRLSPETKRI SVFIQTDKALYRKQEVKRIYTLFSDFXPY  
KTSILNLIKDKPKNLILQOMLSQSDGLSVSKTFOLSHSHILGMSLOVONQDTYOS  
FOYSEVLYPKFEVTLQTPLYCSMNSKNGITAKTYTKPVKGVDTTLFLSPMGK  
KNITTEPKINSFNSPDEKKNYMSNGISELIDSSREPVIITLTVTSVNGI  
SRVSTNVKPKGNDYIIFPDYTTVALKPSLFAVTKVTRADSDQNLLEERNNVIT  
VTRNTEYWSGNSGNQKMEAVQKINLYVPSGTAKVIFLIDSSLEQLAAYFGS  
KSMVAHSLFKSPSKTYIQIKTRDENIKVSGPELVSGNKRLEKLSNVVSGQLVA  
VKONSTMSLTPENSWTPKACAVIYVIEDDBEIIISDVAKIPQALVFKKIKLYMSKV

KAEPSEKVSIRISTVQPDISIVGIVAVDKSVNLMAASNDITMENVHELELYNTGYLIG  
MEMNSFAFOEGLMVLTDANLTKOYIGGVNDNAEYARFMEENGHIVDIDFELGS  
SPVAKHPETWIMLIDTMNGYRIYOEFTVWDSTTSVWAGFVLSBGLGLTTTPV  
ELIAROPPEIIPNIPYSVIRBEEPALEITINNYLKADEEVUJIEKSPILIMTMSNE  
INATGHOQTLIVPSEDGATVULFPIRPTHIGEIPVTVALSPASAVTOMILYKAGI  
EKSYISQSLILDDTDNRLOSITLKTSFSPPTVTVSGSEVQITAGIDVPSINGLASL  
IRMPYGCENQNMNINPANIYILDYITKKOILDNKEALSFMRGXYRELIYOREDG  
SEFARGNDPSSGISTWLSAFVLRCELEADPYIDIDQNVILHRTYTWLKGHOKSGEFWDP  
GRVISEILOGGKSPVTLTAVYIVTSILGTYRKQYVR1"  
BASE COUNT 1108 a 675 c 731 g 1021 t  
ORIGIN  
Alignment Scores:  
Pred. No.: 0 Length: 3535  
Score: 5470.00 Matches: 1062  
Percent Similarity: 100.00% Conservative: 1  
Best Local Similarity: 99.91% Mismatches: 0  
Query Match: 74.44% Indels: 0  
DB: 6 Gaps: 0  
US-10-020-095-4 (1-1428) x AX534942 (1-3535)  
QY 1 MetGlnGlyProProleuLeuThra1a1ahisleuLeuCyvala1CysThra1a1aleu 20  
113 ATGCAAGGGCCACCGCTCTGACCGCCGCCACCTCTGCGTGGACACCGCGCTG 172  
QY 21 AlaValAlaProGlyProArgPheLeuValThra1a1ProGlyIle1a1ArgProGly 40  
173 GCGGTGGCTCCCGGCGCTGGGTTCTGGAGAGGCCACGAGATATACAGGCGGAGA 232  
QY 41 AenValThra1leGlyVala1GluLeuGluHisCysProserGlnVala1Thra1lysa1a 60  
233 AATGAGACTATGGGGGTGGAGCTTCTGGAACACTGCCCTTACACAGTACTGTAGAGCG 292  
QY 61 GluLeuLeuTyrrThra1SerSerleuThra1SerValleuGluAlaGluGlyVala1phe 80  
293 GAGCTGCTCAAGACAGACATCAAACTCACTGCTCTGCTGGAGACAGAGAGACTTTT 352  
QY 81 GluTyraGlySerPheleThreThreThreProserleuProleuansSerAlaaspGlu 100  
353 GAAAAAGGCTCTTTAAGACACTTACTCTTCATACACTTCAACAGTGCAGATAG 412  
QY 101 IleTyrgluLeuArgVala1ThrglyArgThrglna1a1leuPheSeranserThr 120  
413 ATTATGAGCTACGCTTAACCGAGCTACCGAGATGAGATTTATTTCTTAATAGTAC 472  
QY 121 ArgLeuSerPheGluThra1lysa1gileservalpheileglnThra1aspLyala1leuTyrr 140  
473 CGCTTATCATTTGAGACCAAGAGATATCTGCTTCATTTCAAAACAGACAGAGCTTATAC 532  
QY 141 TyrrProLyGlnGluVala1lysa1pheArgIleVala1ThreLeuPheSeraspPheTyrr 160  
533 AAGCCAAAGCAAGAGTAAAGTTGGCATTTGTAACCTTCTCGATTTTAAAGCTTAC 592  
QY 161 TyrrThra1Serleuans1leu1lelysa1aspProlySeransleu1leGlnThrepu 180  
593 AAAACCTTTTAAACATTTCTATTAGAACCCCAATCAATTTATTCACACAGGTG 652  
QY 181 SerGlnGlnSeraspLeuGlyVala1leSerlyThra1pheGlnLeuSerSerHisaspProle 200  
653 TCACAAACAAAGTACCTTGAGATCATTTCCAAAACCTTTTACAGCTATCCCATCAATA 712  
QY 201 LeuGlyaspTyrrPse1leGlnVala1GlnVala1aenaspGlnThra1TyrrGlnSerPheGln 220  
713 CTGGGTGATGCTCTATTCAAGTTCAAGTGAATGACAGACATATTATCAATCAATTTAG 772  
QY 221 ValSerGluTyrrVala1leuProlyaspPheGluVala1Thra1leuGlnThra1ProleuTyrrCysSer 240  
773 GTTTAGATATATGATTTACAAAATTTAAGTGAAGCTTTGACAGACACCATTAATTTGTTCT 832  
QY 241 MetAenSerTyrrHisleuansnGlyThra1leThra1a1aTyrrThra1GlyLyseProVal 260  
833 ATGAATTTCAAGCATTTTAAATGATCATCAAGGCAAGTATACATATGGAGACCAAG 892

QY 261 LysGluAspValThrLeuThrPheLeuProLeuSerPheTrpGlyLysLysLeuAsnIle 280  
 DB 893 AAGAGAGCGTAACGCTTACATTTTACCTTATCTTTGGGAAAGAAAGAAAAATTT 952  
 QY 281 ThrLysThrPheLysIleAsnGlySerAlaAsnPheSerPheAsnAspGluMetLys 300  
 DB 953 ACAAAAACATTTAAGATTAATGAGATCTGCAAACTCTCTTTAATGATGAAGACATGAAA 1012  
 QY 301 AsnValMetAspSerSerAsnGlyLeuSerGlyIleuLeuAspLeuSerSerProGlyPro 320  
 DB 1013 AATGTAATGATTTCTTCAATGAGACTTTCTGAATACTGGATCTATCTTCCCTGGACCA 1072  
 QY 321 ValGluIleLeuThrThrValThrGluSerValThrGlyIleSerArgAsnValSerThr 340  
 DB 1073 GTAGAAATTTTAAACACAGTACAGAAATCAGTTACAGTATTTCAAGAAAGTAAAGCACT 1132  
 QY 341 AsnValPhePheLysGlnHisAspTyrIleIleGluPhePheAspTyrThrThrValLeu 360  
 DB 1133 AATGTTCTTCAAGCAACATGATTAATCATCTGATTTTGAATTAATCACTACTGCTTG 1192  
 QY 361 LysProSerLeuAsnPheThrAlaThrValLysValThrArgAlaAspGlyAsnGluLeu 380  
 DB 1193 AAGCATCTCTCACTTACAGCACTGAGGAACTGAGTCTGCTGATGGCAACCACTG 1252  
 QY 381 ThrLeuGluGluArgArgAsnAsnValValIleThrValThrGlnArgAsnTyrThrGlu 400  
 DB 1253 ACTCTGAAGAAAGAAAGAAATATGTAATCATTAACAGTACAGACAGAAATATATCTGAG 1312  
 QY 401 TyrTrpSerGlySerAsnSerGlyAsnGlnLysMetGluAlaValGlnLysIleAsnTyr 420  
 DB 1313 TACTGAGCGGATCTAACAGTGAATTCAGAAATGCAAGCTGTCAGAAATTAATATAT 1372  
 QY 421 ThrValProGlnSerGlyThrPheLysIleGluPheProIleLeuGluAspSerSerGlu 440  
 DB 1373 ACTGCCCCCAAGGGAATCTTTAAGATTAATTCCTCAATCTCGAGGATTCACAGTAC 1432  
 QY 441 LeuGlnLeuLysAlaTyrPheLeuGlySerLysSerSerMetAlaValHisSerLeuPhe 460  
 DB 1433 CTACAGTTGAAGGCTTATCTCTGTAAGTAAAGTACAGTCAAGTCAATGCTGTT 1492  
 QY 461 LysSerProSerLysThrTyrIleGlnLeuLysThrArgAspGluAsnIleLysValGly 480  
 DB 1493 AAGTCTCTAGTAAGACATACATCACTAAACCAAGACATGAATAATTAAGTGGGA 1552  
 QY 481 SerProPheGluLeuValValSerGlyAsnLysArgLeuLysGluLeuSerTyrMetVal 500  
 DB 1553 TCGCCTTTTGAAGTGTGTGAGTGAAGCAACAGATTTGAAGAGTAAAGCTATATGGA 1612  
 QY 501 ValSerArgGlyGlnLeuValAlaValGlyLysGlnAsnSerThrMetPheSerLeuThr 520  
 DB 1613 GTATCCAGGGGACAGTGTGTGTGTAGAGAAACAAATTCACAAATCTCTTTAACA 1672  
 QY 521 ProGluAsnSerTrpThrProLysAlaCysValIleValTyrTyrIleGluAspAspGly 540  
 DB 1673 CCAAAAATTTCTTGACTCCAAAAGCTGTGAATTTGTATATATTGAAGATGATGGG 1732  
 QY 541 GluIleIleSerAspValLeuLysIleProValGlnLeuValPheLysAsnLysIleLys 560  
 DB 1733 GAAATTAATAGTGTCTTAAATAATCTCTGTCAGCTGTTTAAATAATAATGAATGAAG 1792  
 QY 561 LeuTyrTrpSerLysValLysAlaGluProSerGluLysValSerLeuArgIleSerVal 580  
 DB 1793 CTATATTTGAGTAAGTAAGTAAGTAAGTAAGTAAGTAAGTAAGTAAGTAAGTAAGTAAG 1852  
 QY 581 ThrGlnProAspSerIleValGlyIleValAlaValAspLysSerValAsnLeuMetAsn 600  
 DB 1853 ACACAGCCTGACTCATATGTTGGATTTGACTGTGACAAAGAGTGAATCTGATGAT 1912  
 QY 601 AlaSerAsnAspIleThrMetGluAsnValValHisGluLeuGluLeuTyrAsnThrGly 620  
 DB 1913 GCCTCTATATGATATTACAAAGTAAGTAAGTGTGTCCATGATGGAATTTATTAACAGAGA 1972

QY 621 TyrTyrLeuGlyMetPheMetAsnSerPheAlaValPheGlnGluCysGlyLeuTrpVal 640  
 DB 1973 TATTAATTTAGGATGATTCATGAATTTCTTTGCACTCTTTCAGGAATGTGACTCTGGGA 2032  
 QY 641 LeuThrAspAlaAsnLeuThrLysAspTyrIleAspGlyValTyrAspAsnAlaGluTyr 660  
 DB 2033 TTGACAGATGCAAACTCCACAGAGATTAATTAATGAGTGTATTATACAAATGACAAATAT 2092  
 QY 661 AlaGluArgPheMetGluGluAsnGluLysIleValAspIleHisAspPheSerLeu 680  
 DB 2093 GCTGAGAGGTTTATGAGAGAAATGAAGACATATGTATGATATCTATGACTTTCTTTG 2152  
 QY 681 GlySerSerProHisValArgLysHisPheProGluThrTrpIleTrpLeuAspThrAsn 700  
 DB 2153 GTAGCAGTCCACATGCTCCAAAGCAATTTCCAGACACTTGATTTGGCTAGACACCAAC 2212  
 QY 701 MetGlyTyrArgIleLysGlnGluPheGluValThrValProAspSerIleThrSerTrp 720  
 DB 2213 ATGGGTTACAGATTTTACAAAGATTTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 2272  
 QY 721 ValAlaThrGlyPheValIleSerGluAspLeuGlyLeuGlyLeuThrThrProVal 740  
 DB 2273 GTGGCTACTGGTTTGTGATCTTGAGGACTGGGTCTTGACTTAACACTACTCCAGTG 2332  
 QY 741 GluLeuGlnAlaPheGlnProPhePheIlePheLeuAsnLeuProTyrSerValIleArg 760  
 DB 2333 GAGCTCCAACTCTCCAACTATTTTCAATTTTGAATCTTCCCTACTCTGTTATCAGA 2392  
 QY 761 GlyGluGluPheAlaLeuGluIleThrIlePheAsnTyrLeuLysAspAlaThrGluVal 780  
 DB 2393 GGTAAGAAATTTCTTGGAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2452  
 QY 781 LysValIleIleGluLysSerAspLysPheAspIleLeuMetThrSerSerGluIleAsn 800  
 DB 2453 AAGTATCTATTAAGAAAGTGAAGCAATTTGATTTCTAATGATCTTCAATGAATTAAT 2512  
 QY 801 AlaThrGlyHisGlnGlnThrLeuLeuValProSerGluAspGlyAlaThrValLeuPhe 820  
 DB 2513 GCCACAGGCGCACAGACAGACCTTCTGTGTTCCAGTGAAGATGGGCAACTGTTCTTTT 2572  
 QY 821 ProIleArgProThrHisLeuGlyGluIleProIleThrValThrAlaLeuSerProThr 840  
 DB 2573 CCCATCAGGCGCAACATCTGGAGAAATTCATATCAAGTCAACACTCTTCCACCACT 2632  
 QY 841 AlaSerAspAlaValThrGlnMetIleLeuValLysAlaGluGlyIleGluLysSerTyr 860  
 DB 2633 GCTTCTGATGCTTACCCAGATGATTTTATTAAGAGCTGAAGAAATGAAGAAATCATAT 2692  
 QY 861 SerGlnSerIleLeuLeuAspLeuThrAspAsnArgLeuGlnSerThrLeuLysThrLeu 880  
 DB 2693 TCACATCCATCTTATTAAGCTTGACTGACCAATAGCTACAGATGCCCTGAAGAACTTTG 2752  
 QY 881 SerPheSerPheProProAsnThrValThrGlySerGluArgValGlnIleThrAlaIle 900  
 DB 2753 AGTTTCTCATTTCTCTTATATACAGTCTGGCAGAGAAAGTAAGTCACTGCAAT 2812  
 QY 901 GlyAspValLeuGlyProSerIleAsnGlyLeuAlaSerLeuIleArgMetProTyrGly 920  
 DB 2813 GAGAGATGTTCTTGTCTTCCATCATAGGCTTATGCTTATGATTTGAGTGGCTTATGGC 2872  
 QY 921 CysGlyGluGlnAsnMetIleAsnPheAlaProAsnIleTyrIleLeuAspTyrLeuThr 940  
 DB 2873 TGTGTGAACAGAAACATGAATTAATTTTGTCTCCAAATATTTTACATTTTGTATTCGACT 2932  
 QY 941 LysLysLysGlnLeuThrAspAsnLeuLysGluLysAlaLeuSerPheMetArgGlnGly 960  
 DB 2933 AAAAAGAAACAACTGACAGATTAATTTGAAGAAAGAAAGCTCTTTCATTTATGAGCAAGT 2992  
 QY 961 TyrGlnArgGluLeuLeuTyrGlnArgGluAspGlySerPheSerAlaPheGlyAsnTyr 980  
 DB 2993 TACCAGAGAACTTCTATTCAGAGGAAAGAGGCTTTTCAGTCTTTTGGGAAATAT 3052  
 QY 981 AspProSerGlySerThrTrpLeuSerAlaPheValLeuArgCysPheLeuGluAlaAsp 1000

DB 3053 GACCCCTTGGAGGACCTTGTCAGCTTTGTTTAAAGATGTTCCCTTAAGCCGAT 3112  
QY 1001 ProTyrIleAspIleAspIleAsnValleuHlsArgThrTyrThrTyrleuylsGlyHis 1020  
DB 3113 CCTTACATGATATTCATCATGATATGTTTACACAGACATACACTTGGCTTAAAGACAT 3172  
QY 1021 GlnylSerAsnGlyGlnIleuThrPaeProGlyIArgValIleHisSerGlnleuGlnGly 1040  
DB 3173 CAGAAATCCACGCGTGAATTTGGGATCCAGGAAGATGATTCATGATGAGCTTCAAGCT 3232  
QY 1041 GlnylSerProValThrIleuThrAlaTyrIleValThrSerleuGlnGlyTyrArg 1060  
DB 3233 GGCATTAAGGATCCAGTACACTTACAGCTATATGTAACCTTCTCTCGGAGATATAGA 3292  
QY 1061 LysTyrGln 1063  
DB 3293 AGGTATCAG 3301  
RESULT 9  
AX534944 3535 bp DNA linear PART 22-NOV-2002  
LOCUS AX534944  
DEFINITION Sequence 11 from Patent WO02070696.  
ACCESSION AX534944  
VERSION AX534944.1 GI:25261494  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.  
REFERENCE  
AUTHORS Schuh, A. and Sutherland, R.D.  
TITLE Cid09 nucleic acid molecules polypeptides and methods of use  
JOURNAL Patent: WO 02070696-A 11 12-SEP-2002:  
Schuh, Andre (CA) ; Sutherland, Robert D. (CA)  
FEATURES  
source  
1. .3535  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
113. .3316  
/note="unnamed protein product"  
/codon\_start=1  
/protein\_id="CAD57236.1"  
/db\_xref="GI:25261495"  
/translation="MQGPEPLTAAHLCCVCTAALAAVAPRPLVTAAPGIRGCVTI  
GVELLEHCPQVTAELKLTASNLTVSLAEAGVPEKGSFTTLPLSLPLNSADEIY  
ELRVTRQDEILFNSNSTRLSPEKRIISVFIQTDALYKPKQEVERIVTLPSDPEKY  
KSLNLILKDPKSNLICOQLSQSDLGIVSKTFOLSHPLIDMSIQVQNDQTYOS  
FQVSEYVLPKFEVTLQTPLYCSMNSKHLNGTTAKYTGKPKGAVTLTLPILSPMGK  
KNITKTEKINGSANSPENDEMKVMSNGLSEYLDLSGPEVILITTTVESTGT  
SRNVSTNPFKQHDYIIEFDYTTVLKSLNLTATVATKTRADGNOLTEBRNNVIT  
VQRNYTEWSSNSGNDQMEAVOKINTVPOSQFVKLEPILDESSSELQAAVPLGS  
KSMVAHSLFKPSKTYIQLKTRDENIKVSGPELVSGNKLKELSLVWSRGLVA  
VGKQNSTWESLTPENSWPKACVIVYIEDDEILSDYIKIIVQVLEFKKIKIYMSKY  
KAPEBKSLRISTVOPDSIVGIVVDSVNMASNDITMENVVHLEIKLVTGYIG  
MEMNSFAYROEGILWLDANLTROYIGVYNMAYARPMENSGHYDIDHESLGS  
SPHVRKRPETWIDLDNMGKRIYQEFETVDSITSWATGFIISBDGLATTPV  
ELQAFQPFIFLNLVSIRGEFALEITIFNYLADAEVKYIIEKSKFDLMTTNS  
INATHQOQLLVSESDGATVLPPIRPHLGEIPITVTALSPASADVMILVKAGLI  
EKSYQSILDLTDNRLOSTLKTLSFSPPTVGSERQIATIDVAGSINGASLI  
IRMPYGCGRNNINFAPNITLYLDVLTKKKOLTDNLKEXLASPMRGYORELIVQEDG  
SPSARGNYDRGSTWLSAFLVLCPLADPDIIDQNYLHRTYTNLKGHOKSNGEHWDP  
GHVHSELQGGKSPVTLTAIVYISLGRKTYQVFRF"

DB: 6 Gaps: 0  
US-10-020-095-4 (1-1428) x AX534944 (1-3535)  
QY 1 MetGlnGlyProProleuLeuThrAlaAlaHisleuLeuCyseThrAlaAlaLeu 20  
DB 113 ATGACGGGCCCCACCGCTCTGACCGCCGCCACCTCCCTGCTGTGTGACCGCCCGCTG 172  
QY 21 AlaValAlaProGlyProArgPheLeuValThrAlaProGlyIleIleArgProGly 40  
DB 173 GCCGTGCTCCCGGCTCCGCTGCTTCTGTGTGACGCCCCAGGATATACAGCCCGGAGCA 232  
QY 41 AsnValThrIleGlyValGlnleuGlnIleHisCyProSerGlnValThrValIysAla 60  
DB 223 AATGTGACTATTGGGGTGGAGCTTGGAACACTGCCCTTCAACAGTACTGTGAAGGG 292  
QY 61 GlnleuLeuLysThrAlaSerAsnleuThrValSerValleuGlnAlaGlnGlyValPhe 80  
DB 293 GAGCTGCTCAAGACAGCATCAAACTCACTGCTCTGCTGGAAGCAGAAAGAGCTCTT 352  
QY 81 GlnylSerPheLeuThrleuThrleuProSerleuProleuAsnSerAlaAspGlu 100  
DB 353 GAAAAAGGCTCTTTAAGACACTTACTCTTCATCTACTACTGTAACAGTGCAGATAG 412  
QY 101 IleTyrGlnleuArgValThrGlyArgThrGlnAspGlnIleleuPheSerAsnSerThr 120  
DB 413 ATTTATGAGCTACGCTGTAAACCGAGCTGACCGAGATGAGATTTATCTCTATATGTAAC 472  
QY 121 ArgleuSerPheGluThrLysArgIleSerValPheIleGlnThrAspLysAlaLeuTyr 140  
DB 473 CGCTTATCATTTGAGACCAAGCAAGATATCTGCTTCATCAACACAGACAGAGCTTATAC 532  
QY 141 LysProlGlnGlnValIleValPheArgIleValThrleuPheSerAspPheLysProl 160  
DB 533 AAGCCAAAGCAGAAAGTAAAGTTGCACTGTGTACCTCTTCGATTTTAAAGCTTATAC 592  
QY 161 LysThrSerleuAsnIleleuIleLysAspProlYsserAsnleuIleGlnGlnThrleu 180  
DB 593 AAAACCTCTTAAACATCTCTATTAAGACCCCAATCAATTTGATCAACAGTGTG 652  
QY 181 SerGlnIleSerAspLeuGlyValIleSerLysThrPheGlnleuSerSerHisProlle 200  
DB 653 TCACAAACAAAGGATCTTGAGTCACTTCCAAAACCTTTCAGCTTCCCATCAATA 712  
QY 201 LeuGlyAspTTPSerIleGlnValGlnValAsnAspGlnThrTyrTyrGlnSerPheGln 220  
DB 713 CTGGTGACGTGTCATTAAGGTTCAAGTGAATGACCAAGATATTAATCAATTCATTCAG 772  
QY 221 ValSerGlnTyrValleuProLysPheGlnValThrleuGlnThrProleuTyrCysSer 240  
DB 773 GTTTCAGAAATGTAATTAACAAAATTTAAGTGAATTCGACAGACCAATTAATGTTCT 832  
QY 241 MetAsnSerLysHisleuAsnGlyThrIleThrAlaLysTyrThrTyrGlyLysProVal 260  
DB 833 ATGAATCTTAAGCATTTAAATGATGATCAATCAACGGAAGTATCAATGGGAAGCAATG 892  
QY 261 LysGlyAspValThrleuThrPheleuProleuSerPheTyrGlyLysLysAsnIle 280  
DB 893 AAAGAGAGCGTAAACCTTCACTTTTACCTTTTGGGGGAAAAGAAAATAATTT 952  
QY 281 ThrLysThrPheLysIleAsnGlySerAlaAsnPheSerPheAsnAspGlnIleLys 300  
DB 953 ACAAATAACATTTAATATTAATGATCTCAAACTTCTTTTAAAGATGAAGAGTGAATA 1012  
QY 301 AsnValMetAspSerSerAsnGlyLeuSerGlnTyrleuAspLeuSerSerProGlyPro 320  
DB 1013 AATGTATGATGATTTCTTCAATGAGACTTCTGAAATACCTGGAATCTATCTTCCCTGAGCA 1072  
QY 321 ValGlnIleleuThrThrValThrGlnSerValThrGlyIleSerArgAsnValSerThr 340  
DB 1073 GTAGAAATTTTAAACCAAGTACAGAAACAGTATTCAGATATTTCAAGAAATGTAAGCACT 1132  
QY 341 AsnValPhePheLysGlnHisAspTyrIleIleGlnPhePheAspTyrThrThrValleu 360

## Alignment Scores:

Pred. No.: 0 Length: 3535  
Score: 5461.00 Matches: 1061  
Percent Similarity: 99.91% Conservative: 1  
Best Local Similarity: 99.81% Mismatches: 1  
Query Match: 74.32% Indels: 0



```

Db 1133 AATGTCCTTCAAGCAACATGATTAACATCATGTTTGGATTATACACTGCTCTGG 1192
Qy 361 LysProSerLeuAenPheThrAlaThrValLysValThrArgAlaAspGlyAsnGluLeu 380
Db 1193 AAGCACTCTCAACTTCACAGCCACTGTGAAGGTAACCTGCTGAGGCAACCAACTG 1252
Qy 381 ThrLeuGluGluArgAsnAsnValValIleThrValThrGlnArgAsnTyrThrGln 400
Db 1253 ACTCTTGAGAGAAAGAAATTAATGTAGTCAATAAGTGAACAAGAGAAATTAATCTAG 1312
Qy 401 TyrTrpSerGlySerAsnSerGlyAsnGlnLysMetGluAlaValGlnLysIleAsnTyr 420
Db 1313 TACTGGAGCGGATCTCAACAGTGAATCAGAAATGGAAGCTGTCGAGAAATTAATAT 1372
Qy 421 ThrValProGlnSerGlyThrPheLysIleGlnPheProIleLeuGluAspSerSerGlu 440
Db 1373 ACTGTCCTCCAAAGGGAACCTTTTAAGTTGAATTCCTCAATCTCGAGGATTCAGTGA 1432
Qy 441 LeuGlnLeuLysAlaTyrPheLeuGlySerLysSerSerMetAlaValHisSerLeuPhe 460
Db 1433 CTACAGTTGAAGGCTTAATTCCTTGTAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 1492
Qy 461 LysSerProSerLysThrTyrIleGlnLeuLysThrArgAspGluAsnIleLysValGly 480
Db 1493 AAGTCTCTAGTAAAGACATACATCACTAAACAAAGAGATGAATAATTAAGGTGGA 1552
Qy 481 SerProPheGluLeuValIleSerGlyAsnLysArgLeuLysGluLeuSerTyrMetVal 500
Db 1553 TCGCTTTTGAGTGGTGGTGTAGTGAACAACAACATGAAGAGTTAAGCTAATAGTGA 1612
Qy 501 ValSerArgGlyGlnLeuValAlaValGlyLysGlnAsnSerThrMetPheSerLeuThr 520
Db 1613 GTATCCAGGGGACAGTTGGTGGTGTGAAGAAACAATTCACACATGTTCTTTAACA 1672
Qy 521 ProGlnAsnSerTyrPheProLysValCysValIleValTyrTyrIleGlnAspAspGly 540
Db 1673 CCAGAAATTCCTTGACCTCCAAAGCCCTGTGAATGTGTATTAATGAAGATGATCGG 1732
Qy 541 GluIleIleSerAspValLeuLysIleProValGlnLeuValPheLysAsnLysIleLys 560
Db 1733 GAAATTAATAGTGAATGTTCTAAATAATTCCTGTCTGCTGTTTAAATAATGAATGA 1792
Qy 561 LeuTyrTrpSerLysValLysAlaGluProSerGluLysValSerLeuArgIleSerVal 580
Db 1793 CTAATTTGAGACTAAAGTGAAGCTGAACCACTGAAGAGTCTCTTTAGATCTCTGTG 1852
Qy 581 ThrGlnProAspSerIleValGlyIleValAlaValAspLysSerValAsnLeuMetAsn 600
Db 1853 ACACAGCTGACCTCAATGTTGGGATGTGACTGTTGCAAAAGTGAATCTGATGAT 1912
Qy 601 AlaSerAsnAspIleThrMetGluAsnValValHisGluLeuGluLeuTyrAsnThrGly 620
Db 1913 GCCCTTAATGATTAATTAATGAAATGTGTGTCATGATGTGAACCTTTAATGAACAGA 1972
Qy 621 TyrTyrLeuGlyMetPheMetAsnSerPheAlaValPheGlnGluCysGlyLeuTyrVal 640
Db 1973 TATTAATTTAGCACTTCATGATATCTTTTGCAGTCTTTTCAGAAATGTGGAGTGA 2032
Qy 641 LeuThrAspAlaAsnLeuThrLysAspTyrIleAspGlyValTyrAspAsnAlaGluTyr 660
Db 2033 TTGACAGATGCAAACTCAGAGAGATTAATGATGTGTTTGAACAAATGAGAAATAT 2092
Qy 661 AlaGluArgPheMetGluGluLysGlnGlyHisIleValAspIleHisAspPheSerLeu 680
Db 2093 GCTAGAGAGTTTATGAGAGAAATGAAGACATATGTAGATATTCAGATCTTTTCTTGG 2152
Qy 681 GlySerSerProHisValArgLysIlePheProGluThrTyrIleTyrPheAspThrAsn 700
Db 2153 GGTAGCACTCCACATGTCGAAAGCAATTTTCAGAGACTTGGATTTGGCTTGAACCAAC 2212
Qy 701 MetGlyTyrArgIleTyrGlnGluPheGluValThrValProAspSerIleThrSerTyr 720

```

---

```

Db 2213 ATGGGTTCCAGATTATCAAGAAATTTGAAGTAACGTACGTATCTATCACTTCTGG 2272
Qy 721 ValAlaThrGlyPheValIleSerGluAspLeuGlyLeuGlyLeuThrThrProVal 740
Db 2273 GTGGCTACTGTTTGTGATCTCTGAGAGACTGGGCTTGGACTTAACACTACTCCAGTG 2332
Qy 741 GluLeuGlnAlaPheGlnProPhePheIlePheLeuAsnLeuProTyrSerValIleArg 760
Db 2333 GAGCTCCAGACCTTCCAAACATTTTTCATTTTGTGAATCTTCCCTACTCTGTTATAGA 2392
Qy 761 GlyGluGluPheAlaLeuGluIleThrIlePheAsnTyrLeuLysAspAlaThrGluVal 780
Db 2393 GGTGAATAATTCCTTGGAAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2452
Qy 781 LysValIleIleGluLysSerAspLysPheAspIleLeuMetThrSerSerGluIleAsn 800
Db 2453 AAGGTATCTATTGAGAAAGTGAACAAATTTGATTAATTTATGACTTCAATGAATTAAT 2512
Qy 801 AlaThrGlyHisGlnGlnThrLeuLeuValProSerGluAspGlyAlaThrValLeuPhe 820
Db 2513 GCCACAGGCCACAGCAGACCTTCTGTTCCAGTGAAGATGGGGCACTGTTCTTTT 2572
Qy 821 ProIleArgProThrHisLeuGlyGluIleProIleThrValThrAlaLeuSerProThr 840
Db 2573 CCATCAGGCCAACACATCTGGAGAGAAATTCATCACTCACTCAGCTTTTCAACCCT 2632
Qy 841 AlaSerAspAlaValThrGlnMetIleLeuValIleValAlaGluGlyIleGluLysSerTyr 860
Db 2633 GCTTCTGATCTGTCAACCAAGATGATTTTGTAAAGCTTAAGAAATGAATAATCATAT 2692
Qy 861 SerGlnSerIleLeuLeuAspLeuThrAspAsnArgLeuGlnSerThrLeuLysThrLeu 880
Db 2693 TCACATTCATCTTATTAAGACTGACTGACATATAGCTACAGATCACTGAAACCTTGG 2752
Qy 881 SerPheSerPheProProAsnThrValThrGlySerGluArgValGlnIleThrAlaIle 900
Db 2753 AGTTTCATATTCCTCTAATTAACAGTACGTGAGTGAAGAAAGTTCAGATCACTGCAATT 2812
Qy 901 GlyAspValLeuGlyProSerIleAsnGlyLeuAlaSerLeuIleArgMetProTyrGly 920
Db 2813 GGAATGTTCTTGTCTCTTCATCATATGCTTGAAGCTTCAATGCTTATGCTTATGGC 2872
Qy 921 CysGlyGlnGlnAsnMetIleAsnPheAlaProAsnIleTyrIleLeuAspTyrLeuThr 940
Db 2873 TGTGTAAACAGAAACATGAATTAATTTGCTCCAAATATTAATTTGATTAATGACT 2932
Qy 941 LysLysLysGlnLeuThrAspAsnLeuLysGluLysValLeuSerPheMetArgGlnGly 960
Db 2933 AAAAAGAAACAACTGACAGATTAATTTGAAGAAAGAAAGCTCTTCAATTAAGAGCAAGT 2992
Qy 961 TyrGlnArgGluLeuLeuTyrGlnArgGluAspGlySerPheSerAlaPheGlyAsnTyr 980
Db 2993 TACCAGAGAAACCTCTCTATCAAGAGGAAAGAGGCTCTTCACTGCTTGTGGAAATTA 3052
Qy 981 AspProSerGlySerThrTyrPheLeuSerAlaPheValLeuArgCysPheLeuGluAlaAsp 1000
Db 3053 GAOCCTCTGGAGCACTTGTGTTGCTGACGCTTTGTTTAAGATGTTTCTTGAAGCCGAT 3112
Qy 1001 ProTyrIleAspIleAspGlnAsnValLeuHisArgThrTyrThrTyrPheLysGlyHis 1020
Db 3113 CTTTACATTAATTAATGATTAATGAAATGTGTTACACAGAAACATCACTTGGCTTAAGGCA 3172
Qy 1021 GlnLysSerAsnGlyGluPheThrAspProGlyArgValIleHisSerGluLeuGlnGly 1040
Db 3173 CAGAAATCCAAAGTGAATTTGGATCCAGAAAGTGAATTAATTAATTAATTAATTAAT 3232
Qy 1041 GlyAsnLysSerProValThrLeuThrAlaTyrIleValThrSerLeuLeuGlyTyrArg 1060
Db 3233 GGCATTAATTAAGTCCAGTAACATTAACGCTAATTTATTAATTTCTCTCCGGGATATAGA 3292
Qy 1061 LysTyrGln 1063
Db 3293 AAGTATCAG 3301

```

[illegible]



Db 1360 TCAGAAATGCGAACCAATGCTGTTTTCAGCAACATGATTAATCATTCATTGAGATTTT 1419  
 Qy AspIYrThrValIleuLysProSerIleuAenPheThrAlaThrValIleValIleThrArg 374  
 Db 1420 GATTATATGACTGCTTGAACCAATCTCTCAATTCACAGCCCATCTGGAAGTCAAGCCGC 1479  
 Qy 375 AlaAepGIyaenGIuLeuThrLeuGIuGIuArgArgAenValIleIleThrValIleThr 394  
 Db 1480 TCTGATGGCAATCAACTACTCTGGAAGAAATGAAATGATTTGGTACAGTAGTGACC 1539  
 Qy 395 GlnArg-----AenIYrThrGIuYrTrpSerGIySerAenSerGIyAenGIuLysMet 412  
 Db 1540 CAGCGGAAAAATTAACCACTCTGAG-----AGTCAAGAGATCAGAGAAATG 1584  
 Qy 413 GluAlaValGIuLysIleAenIYrThrValIleProGIuSerGIyThrPheLysIleGIuPhe 432  
 Db 1585 GATTATATCCAGACCGTAAATTAATTAATCCCAAAATGGAATCATTAAGATTGAATTC 1644  
 Qy 433 ProIleLeuGIuAepSerSerGIuLeuGIuLysAlaIYrPheLeuGIySerLysSer 452  
 Db 1645 CCAAGTCAATGCTCATTTCCGGCGAGCTACAACTGAAGGCTTATTTCTTGATGTACAAGT 1704  
 Qy 453 SerMetAlaValIleSerLeuPheLysSerProSerIYrTrpIleGIuLysThr 472  
 Db 1705 TCTGTGACAGTCCACAGTATGTTCACTCTCTAGTAAAGCATCCACCTTAAACC 1764  
 Qy 473 ArgAepGIuAenIleLysValIleSerProPheGIuLeuValIleSerGIyAenLysArg 492  
 Db 1765 AGAATGATTAATTAATTAAGTGGGCTCACTTTGATTTGATGTAGTGGCAACAGACAA 1824  
 Qy 493 LeuLysGIuLeuSerIYrMetValIleSerArgIleGIuLeuValIleValIleLysGIu 512  
 Db 1825 TTCAAGACCTTAATGATATGATATCAAAAGGACGTTGGTGGCTGACAGGCAAAACA 1884  
 Qy 513 AsnSerThrMetPheSerLeuThrProGIuAenSerIYrProLysAlaLysValIle 532  
 Db 1885 AGTTCAGAACTTTCTCTTAACACCAAGGCTTCTGGGCTCCAAAGGCTGTATATC 1944  
 Qy 533 ValIYrTrpIleGIuAepAepGIyGIuIleIleSerAepValIleLysIleProValIle 552  
 Db 1945 GCATATTAATTAATGAGAGATGGGAAATTAATTAATGATATTTCAAAATATCCCGTTACG 2004  
 Qy 553 LeuValPheLysAenLysIleLysLeuYrTrpSerLysValIleValIleProSerGIu 572  
 Db 2005 CTGTGTTTGAATAATGATTAAGTACTGTTGAGTAACTCAAGTCAAGTCAAGTCAAT 2064  
 Qy 573 LysValSerLeuAenLysIleSerValIleThrGIuProAepSerIleValIleValIle 592  
 Db 2065 AAGGCTCTCTCAGAGATCTCTGCGACACAGTCTGACCTCCGTGGGGATTTGTAGCTGT 2124  
 Qy 593 AspLysSerValAenLysMetAenLysAenAepIleThrMetGIuAenValIleHis 612  
 Db 2125 GACAAAGTGTACATGATGATGAAATCCCAACAGTATTAACAATGAGAAACCATGTCAT 2184  
 Qy 613 GluLeuGIuLeuYrAenThrGIyTrpIYrLeuGIuLysMetPheMetAenSerPheAlaValI 632  
 Db 2185 GAACTGGAATCTTAAACAGATATTAATTAAGCATGTTCAAGACTCTTCTGCTGTA 2244  
 Qy 633 PheGIuGIuLysGIyLeuTrpValIleuThrAspAlaAenLeuThrLysAepTrpIleAep 652  
 Db 2245 TTTCAGAAATGCTGCTCTGGGTGTGACAGATCAACCTTATTAATGATGATGATGAT 2304  
 Qy 653 GluValIYrAepAenAlaGIuYrAlaGIuArgPheMetGIuGIuAenGIuLysIle 672  
 Db 2305 GAATCTACAGTACTGGAAGATTAATTCGAAGCTTTTCAGAGGAAAAATGAGCAAACTG 2364  
 Qy 673 ValAepIleHisAepPheSerLeuGIySerSerProHisValAenLysHisAepProGIu 692  
 Db 2365 GTTGATTTTGAAGATGCTTCTTCACTTAACAATGATACATGTCAGAGAAATTTCCAGAA 2424  
 Qy 693 ThrTrpIleTrpLeuAepThrAenMetGIyTrpArgIleYrGIuGIuPheGIuValIleThr 712  
 Db 2425 ACATGATTTGGCTAAGACGCTCATGAGATCTTAAGATTTTGAAGATTTGAGTACT 2484

Qy 713 ValProAepSerIleThrSerTrpValAlaThrGIyPheValIleSerGIuAepLeuGIy 732  
 Db 2485 GTACCTGACTCCATCACTCTTGGGTGGCTTGGGCTTTTCTCATCTCAGAAACCTGGGT 2544  
 Qy 733 LeuGIyLeuThrThrProValGIuLeuGIuAlaPheGIuProPheIlePheLeu 752  
 Db 2545 TTGGGTAAACAACCGTTCCAGACAGTGCAGACCTTCCAAACGTTTTTCTCTTCTCTG 2604  
 Qy 753 AenLeuProIYrSerValIleArgGIyGIuGIuPheAlaLeuGIuIleThrIlePheAen 772  
 Db 2605 AATCTTCCGATCTGTATCAAGAGTGAAGATTTGCTTGAAGATTCATGTCAT 2664  
 Qy 773 TrpIleLysAepAlaThrGIuValIleIleGIuLysSerAepLysPheAepIle 792  
 Db 2665 TATTTGAAGATTAACAATTAAGTTGATATCTCATGAGGAAATGACACTTGTGATATT 2724  
 Qy 793 LeuMetThrSerGIuIleAenAlaThrGIyHisGIuGIuThrLeuValProSer 812  
 Db 2725 TTATATGCTTCAATATGACAAATGCGACCATTAATCCGGAAGACTGTTCAAGTTCCAGC 2784  
 Qy 813 GluAepGIyAlaThrValIleuPheProIleArgProThrHisIleuGIyGIuIleProIle 832  
 Db 2785 GATTAATGGGTCACTCTGTTTCCATCAAGCCCAAGCTTGGAGAGATTTCCATC 2844  
 Qy 833 ThrValThrAlaLeuSerProThrAlaSerAepAlaValIleThrIleMetIleLeuValLys 852  
 Db 2845 ACAGTACGGCTCCCTCAACCACTGCTGATGCTGTCAACCAAGCAATTTGATGAAG 2904  
 Qy 853 AlaGIuGIyIleGIuLysSerIYrSerGIuSerIleLeuLeuAepLeuThrAepAenArg 872  
 Db 2905 CCGAAGGAGATTAAGAAATGATGATTAATGCAATCTGTCTTATGATCTGACAGATAGCAC 2964  
 Qy 873 LeuGIuSerThrLeuLysThrLeuSerPheSerPheProAenThrValIleThrGIySer 892  
 Db 2965 GTGAAGATTAAGACATTAAGATTTCTCTTCTCTCAATGAGTCAATGGCAGT 3024  
 Qy 893 GluArgValGIuIleThrAlaIleGIyAepValLeuGIyProSerIleAenGIyLeuAla 912  
 Db 3025 GAAGAAGTCAATACAGCAAGCAATGGAAGATATCTGTTCCCTCAATCAATGACTATCT 3084  
 Qy 913 SerLeuIleArgMetProIYrGIyCysGIyGIuGIuAenMetIleAenPheAlaProAen 932  
 Db 3085 TCACATGCTCGAGTGGCTTATGATGATGATGATGATGATGATGATGATGATGATGAT 3144  
 Qy 933 IleTrpIleLeuAepTrpIleuThrLysLysLysGIuLeuThrAepAenLysGIuLys 952  
 Db 3145 ATTACATTTTGAATTAATGACTTAACAGAAACAGTGCAGTTAATTTAAAGAAAG 3204  
 Qy 953 AlaLeuSerPheMetArgGIuGIyTrpGIuArgGIuLeuLeuYrGIuArgGIuAepGIy 972  
 Db 3205 GCCCTTCAATTAATGAGCAAGTTACCAAGGAGACTTCTTAATCAAGAGGAAATGGC 3264  
 Qy 973 SerPheSerAlaPheGIyAenIYrAepProSerGIySerThrTrpLeuSerAlaPheVal 992  
 Db 3265 TCCCTTACGCTTTTGGGACATTAATGACTTCTTGGAGACCTTGGCTGTGACGATTTGTT 3324  
 Qy 993 LeuArgCysPheLeuGIuAlaAepProIYrIleAepIleAepIleAenValIleuHisArg 1012  
 Db 3325 TTAAAGATGCTTTTGAAGCTGATTAATTAAGATTAATGATGATGATGATGATGATGAT 3384  
 Qy 1013 ThrTrpThrTrpLeuLysGIyHisGIuLysSerAepGIyGIuPheTrpAepProGIyArg 1032  
 Db 3385 ACATTAATCTTGGCTTAATGACATTAAGAAATTAAGATGATTAATTTGGAGCCAGAAAG 3444  
 Qy 1033 ValIleHisSerGIuLeuGIuGIyAenLysSerProValIleThrLeuThrAlaIYrIle 1052  
 Db 3445 GTGATTCAGAGTAACTTAAGGTGGACCAAAAGCCAGTAACTTAAGGCTTAATTT 3504  
 Qy 1053 ValThrSerLeuLeuGIyTrpArgGIyTrpGIuProAenIleAepValGIuGIuSerIle 1072  
 Db 3505 GTGACTTCTGTCTGGGATTAACAAAGATTAACCTTAATTAATGATGATTAACAGTCAATC 3564

QY 1073 HApheLeuGluSerGluPheSerArgGlyLeSerAspAsnThrThrLeuAlaLeuLe 1092  
 DB 3565 AAGTTTTGTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3624  
 QY 1093 ThrTyAlaLeuSerSerValGlySerProValAlaGlyAlaLeuMetLeuThr 1112  
 DB 3625 TCTTAATGCTGCTGCAAGCGGAGCCCTTAAGACAGAGAGGCTTTGAATCTGTGATG 3684  
 QY 1113 TPATGAlaGluGlnGluGlyGlyMetGlnPheThrValSerSerGluSerLysLeuSer 1132  
 DB 3685 CAGCATCAAGAAAGAGAGAGACACAGCTTCTGTGTGTCATCAGAGACTGGCTCTCT 3744  
 QY 1133 AspSerTrpGlnProArgSerSerLeuAspIleGlnValAlaAlaValAlaLeuSerHis 1152  
 DB 3745 GGGTCTCTGCGAGCGCTCCGCTCCGTCGCAATCAAGAAATCTGATCACTGCTGCGCAC 3804  
 QY 1153 PheLeuGlnPheGlnThrSerGluGlyIleProIleMetArgThrLeuSerArgGlnArg 1172  
 DB 3805 ACTCTGCAC---CATGTCTGTAGGAGAAATCCGGTTATAGAGTGCTCATCCAGCAAGA 3861  
 QY 1173 AasnSerLeuGlyGlyPheAlaSerThrGlnAspThrThrValAlaLeuLysAlaLeuSer 1192  
 DB 3862 AACAGCTCGAGAGTTTGTATTCACACTCAGAGACTGTGTGCTGAGAGCATTAATCT 3921  
 QY 1193 GluPheAlaAlaLeuMetAsnThrGluArgThrAsnIleGlnValThrValThrGlyPro 1212  
 DB 3922 GAATTTTCAGCCCTTGTGCAAGAAATACAGATATCACTCACTGACCGTCACAGGGCCC 3981  
 QY 1213 SerSerPro----- 1215  
 DB 3982 GGCATCCCGAATCCATACACTCCGAATGACTCCAGAACCTTTCTCTTCCACAG 4041  
 QY 1216 SerProLeuAlaValAlaGlnProMetAlaValAsnIleSerAlaAsnGlyPheGlyPhe 1235  
 DB 4042 GAAGACTTCAGCGCTAGATCCCATCCACTCAGCTTAATGTTTCTGCGGATGCTCGGATTT 4101  
 QY 1236 AlaIleCysGlnLeuAsnValValTyraAsnValLysAlaSerGlySerSerArgArgArg 1255  
 DB 4102 GCTATTTGCCAGCTTAATGTTGCTATTAAGTGAAGGTTTCAAGGTTCTTCTAAGAACAG 4161  
 QY 1256 ArgSerIleGlnAsnGlnGluAlaPheAspLeuAspValAlaValLysGluAsnLysAsp 1275  
 DB 4162 AATATCACCGAAATCTGAAGATTTTGTATTAGAGTCAATGTG---AATATGAGGAGC 4218  
 QY 1276 AspLeuAsnHisValAspLeuAsnValCysThrSerPheSerGlyProGlyArgSerGly 1295  
 DB 4219 GACATTAGCACCTGAATCTGAATGTGTGCAAGAGTCACTTGAGTTTCAGAGAGGACAGGC 4278  
 QY 1296 MetAlaLeuMetGluValAsnLeuLeuSerGlyPheMetValProSerGluAlaIleSer 1315  
 DB 4279 ATGGTCCCTTATGGAAGTGAACCTTCTCAATGCTTAAAGTCAATCTTCAATTCATTCCT 4338  
 QY 1316 LeuSerGluThrValLysValGlyValGlyTyraAspHisGlyLysLeuAsnLeuTyraAsp 1335  
 DB 4339 CTGAGTGAACCTCTGAAGAAAGTGAATGTATTAACGGGAAACTCAACTTTATTTAGAT 4398  
 QY 1336 SerValAsnGluThrGlnPheCysValAsnIleProAlaValArgAsnPheLysValSer 1355  
 DB 4399 TCTGTAATGAATCCAGTTTGTGTAATATTCCTCACTGAGAGACTCAAGAAAGTTTCA 4458  
 QY 1356 AsnThrGlnAspAlaSerValSerIleValAspTyraTyraGluProArgArgGlnAlaVal 1375  
 DB 4459 AATATTCAGATGATGTTCACTATCTGTATGACTACTATGAACCAAGAGACAGAGAGAG 4518  
 QY 1376 ArgSerTyraAsnSerGluValLysLeuSerSerCysAspLeuCysSerAspValGlnGly 1395  
 DB 4519 CCAAGGTACAAACACCAAGGTGAAGCTGTCTCATGTATCACTGTCACAGCACAC---AAC 4575  
 QY 1396 CysArgProCysGluAspGlyLysSerGlySerHisHisSerSerValIlePheIle 1415  
 DB 4576 TGCAGTACACACAGGAGACGAGACGAGCTCCCTTGAAGCTTTCGAGCTTCTGCTGC 4635  
 QY 1416 PheCysPheLysLeuLeuTyraPheMetGlu 1425

Db 4636 TTTTGCTCCGCTCTCTGTGATCTTGCGCA 4665  
 RESULT 11  
 LOCUS BC052443 5644 bp mRNA linear ROD 19-MAY-2003  
 DEFINITION Mus musculus cDNA clone MGC:63463 IMAGE:6406841, complete cds.  
 ACCESSION BC052443  
 VERSION BC052443.1 GI:30851465  
 KEYWORDS MGC.  
 SOURCE Mus musculus (house mouse)  
 ORGANISM Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 1 (bases 1 to 5644)  
 REFERENCE  
 AUTHORS Strusberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,  
 Klausner,R.D., Collins,F.S., Wagner,K.H., Sherman,C.M., Schuler,G.D.,  
 Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,  
 Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,  
 Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,  
 Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,  
 Scheetz,T.E., Brownstein,M.J., Uedini,T.B., Toshiyuki,S.,  
 Carninci,P., Prange,C., Kana,S.S., Loquellano,N.A., Peters,G.J.,  
 Abramson,R.D., Miallady,S.J., Bosak,S.A., McEwan,P.J.,  
 McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,  
 Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hultk,S.W.,  
 Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,  
 Fahy,J., Helton,E., Kettelman,M., Madan,A., Young,A.C., Shevchenko,Y.,  
 Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,  
 Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,  
 Butlerfield,Y.S., Krzywinski,M.I., Skalka,U., Smillius,D.E.,  
 Scherch,A., Schin,J.E., Jones,S.J. and Matre,M.A.  
 Generation and initial analysis of more than 15,000 full-length  
 human and mouse cDNA sequences  
 Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)  
 12477932  
 2 (bases 1 to 5644)  
 JOURNAL Strausberg,R.  
 MEDLINE Direct Submission  
 PUBMED Submitted (15-MAY-2003) National Institutes of Health, Mammalian  
 REFERENCE Gene Collection (MGC), Cancer Genomics Office, National Cancer  
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,  
 USA  
 REMARK NIH-MGC Project URL: <http://mgc.nci.nih.gov>  
 COMMENT Contact: MGC help desk  
 Email: [cgapbs-remail.nih.gov](mailto:cgapbs-remail.nih.gov)  
 Tissue Procurement: Dr. Jim Ian, University of Iowa  
 CDNA Library Preparation: M. Bento Soares, University of Iowa  
 DNA Sequencing by: University of Iowa, Dr. M. Bento Soares and Dr.  
 Thomas L. Casavant.  
 Web site: <http://genome.uiowa.edu>  
 Contact: [bento-soares@iowa.edu](mailto:bento-soares@iowa.edu); [tom-casavant@iowa.edu](mailto:tom-casavant@iowa.edu)  
 Bonaldo,M.F., Akabogu,I., Bair,T., Bair,J., Crouch,K., Davis,A.,  
 Fiehler,K., Keppel,C., Kucaba,T., Lebeck,M., Melo,A., Schaefer,K.,  
 Scheetz,T., Smith,C., Shiri,E., Tack,D., Trout,K., Walters,J.,  
 Casavant,T., Soares,M.B.  
 Clone distribution: MGC clone distribution information can be found  
 through the I.M.A.G.E. Consortium/LINL at: <http://image.lnln.gov>  
 Series: Place: Row: Column: 0.  
 FEATURES  
 source Location/Qualifiers  
 1..5644  
 /organism="Mus musculus"  
 /mol\_type="mRNA"  
 /strain="C57BL/6"  
 /db\_xref="taxon:10090"  
 /clone="MGC:63463 IMAGE:6406841"  
 /tissue\_type="Brain, enriched mouse brain 12.5dp"  
 /clone\_id="NIH\_BMAP\_F00"  
 /lab\_host="DHI0B"

```

CDS      /note="Vector: pYX-ASC"
          105. .4433

```

```

Db      1587  TTCAAGAGACTTAAGCTATATGTAATACAAAGGACAGTGGCTGCAGCAACAA 1646
Qy      513   AsnSerThrMetPheSerLeuThrProGluAsnSerTrpThrProValAlaCysValIle 532
Db      1647  AGTTCAAGAACTTCTCTTTACACACAGAGCTTCTGGGCTCCAAAGGCTGTATAATC 1706
Qy      533   ValTyrTyrIleGluAspArgIleIleIleSerAspValLeuValIleProValGln 552
Db      1707  GCATATTATATTCAGAGATGGGAAATTATAAGATATCTTAATAATCCCCGTTGAG 1766
Qy      553   LeuValPheLeuAsnValIleValLeuTyrTrpSerLeuValIleGluProSerGlu 572
Db      1767  CTGTGTTTGAATAATAGTAAGTAAGCTGTTTGAAGTAACCTACAGTCAAGCATCTGAT 1826
Qy      573   LysValSerLeuArgIleSerValThrGlnProAspSerIleValGlyIleValAlaVal 592
Db      1827  AAGTCTCTCTCGAGATCTTCGCAACAGTCTGACTCCCTGGTGGGAGTGTAGCTGTT 1886
Qy      593   AspLysSerValAsnLeuMetAsnAlaSerAsnAspIleThrMetGluAsnValIleHis 612
Db      1887  GACAAAGGTAACTTGAATGAGAAATCTCCACAGATTATTCATGAGAAACCATGCTCAT 1946
Qy      613   GluLeuGluLeuTyrAsnThrGlyTyrTyrLeuGlyMetPheMetAsnSerPheAlaVal 632
Db      1947  GAAGTGAACCTTATTAACACAGAAATATTACTAGGACATGTCATGAATCTTCGCTGTA 2006
Qy      633   PheGlnGluCysGlyLeuTyrValLeuThrAspAlaAsnLeuThrLysAspTyrIleAsp 652
Db      2007  TTTCAGAAATGTGGTCTCTGGGTTGTGACAGATGCAACCTTATTAAGATTAAGCATTTGAT 2066
Qy      653   GlyValTyrAspAsnAlaGluTyrAlaGluArgPheMetGluGluAsnGlyHisIle 672
Db      2067  GAAGTGTAGATACGTGAAGATTTCTGAACGCTTTCAGAGAGAAATAGAGCAACCTG 2126
Qy      673   ValAspIleHisAspPheSerLeuGlySerSerProHisValArgLysHisPheProGlu 692
Db      2127  GTTGATTTTGAAGATCTTCTTCAGTTAAACATGTACATGTCCAGAAATTTTCCAGAA 2186
Qy      693   ThrTrpIleTrpLeuAspThrAsnMetGlyTyrArgIleTyrGlnGluPheGluValThr 712
Db      2187  ACATGTGATTTGGTACACCGCTACATGAGATCATTAAGATTTATGAAATTTGAAAGTTACT 2246
Qy      713   ValProAspSerIleThrSerTrpValAlaThrGlyPheValIleSerGluAspLeuGly 732
Db      2247  GTACCTGACTCCATCTCTTGGGTGGCTTGGCTTTGTATCTCAAGAAAGCTGGGT 2306
Qy      733   LeuGlyLeuThrThrThrProValGluLeuGlnAlaPheGlnProPhePheIlePheLeu 752
Db      2307  TTGGGTTTAACAACCGTTCAGAGAGCTGCMAAGCTTCACACCGTTTTCCTTTTCCG 2366
Qy      753   AsnLeuProTyrSerValIleArgIleGluGluPheAlaLeuGluIleThrIlePheAsn 772
Db      2367  AATCTTCCGTAACCTCTGTATCAAGGTGAAGTTTGTGTAAGTATCATCTCCAT 2426
Qy      773   TyrLeuLysAspAlaThrGluValLysValIleIleGluLysSerAspLysPheAspIle 792
Db      2427  TATTTAAAGATTAACAATTAAAGTTGTGATCATATTAAGAAAGTACAGCTTTGATATT 2486
Qy      793   LeuMetThrSerSerGluIleAsnAlaThrGlyHisGlnGlnThrLeuLeuValProSer 812
Db      2487  TTATATGACTTCAATGACACCAATGACCATATACCGGAAGCTTTCAAGTTCCAG 2546
Qy      813   GluAspGlyAlaThrValLeuPheProIleArgProThrHisLeuGlyGluIleProIle 832
Db      2547  GATTAAGGGGTCACTTGTGTTTCCATCAAGCCACCACTGTGGAGAGATTTCCCATC 2606
Qy      833   ThrValThrAlaLeuSerProThrAlaSerAspAlaValThrGlnMetIleLeuValLys 852
Db      2607  AAGTATACGGCTCCACCACTGCTGTGATGCTGTACCCAGCAATTTGATGAAGA 2666
Qy      853   AlaGluGlyIleGluLysSerTyrSerGlnSerIleLeuLeuAspLeuThrAspAsnArg 872

```

```

Db      2667  CCTGAAGGAGATGAAGAAAGTCGATTGCAAAATCTGTATTGATTCAGATGACATGCAAC 2726
Qy      873   LeuGlnSerThrLeuLysThrLeuSerPheSerPheProProAsnThrValThrGlySer 892
Db      2727  GTAGAAAGTAAACACACAATCTATGAGATTCTCTTCTCTCTGATACGATGCTATTCGAGT 2786
Qy      893   GluArgValGlnIleThrAlaIleGlyAspValLeuGlyProSerIleAsnGlyLeuAla 912
Db      2787  GAAAGGTTCAATCAACAGCAATTCAGATATCTTGGTTCCTCATCAATGCTTATCT 2846
Qy      913   SerLeuIleArgMetProTyrGlyCysGlyGluGluAsnMetIleAsnPheAlaProAsn 932
Db      2847  TCACGTAGCTCCGATCTCTTATGATGATGTGTGAACGAACATGATATTAATTTGCTCCAAT 2906
Qy      933   IleTyrIleLeuAspTyrLeuThrLysValIleGluLeuThrAspAsnLeuLysGlyLys 952
Db      2907  ATTTACATTTTGATTTATCTGCTTAAACAGAAACGCTGACGTTAAATTTAAAGAAAG 2966
Qy      953   AlaLeuSerPheMetArgGlnGlyTyrGlnArgIleLeuLeuTyrGlnArgGluAspGly 972
Db      2967  GCCCTTCAATATGAGGCAAGTTACCAAGAGAGCTTCTCTATCAGAGGAGAAATGCG 3026
Qy      973   SerPheSerAlaPheGlyAsnTyrAspProSerGlySerThrTrpLeuSerAlaPheVal 992
Db      3027  TCTTTCAGTGTCTTTGGGACATTCATGACTCTTCGGAGACACTTGCTGTACAGATTGTT 3086
Qy      993   LeuArgCysPheLeuGluAlaAspProTyrIleAspIleAspGluAsnValLeuHisArg 1012
Db      3087  TTAAAGATCTTCTTGAAAGCTGATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3146
Qy      1013  ThrTyrThrTrpLeuLysGlyHisIleGlnLysSerAsnGlyGluPheTrpAspProGlyArg 1032
Db      3147  ACATATATCTGCTTAAATGACATTAAGAAATTCATATGATGATTTTGGAGCCAGGAAGA 3206
Qy      1033  ValIleHisSerGluLeuGlnGlyValAsnLysSerProValThrLeuThrAlaTyrIle 1052
Db      3207  GTGATTTCAAGAGACTTCAAGGTGGCACAAAGCCAGTAACCTTAACCGGCTTATATT 3266
Qy      1053  ValThrSerLeuLeuGlyTyrArgLysTyrGlnProAsnIleAspValGlnGluSerIle 1072
Db      3267  GTGACTTCTGCTCTGGGATACAAAGATATGAGCTTAATATGATGATGATGATGATGATGAT 3326
Qy      1073  HisPheLeuGluSerGluPheSerArgGlyIleSerAspAsnTyrThrLeuAlaLeuVal 1092
Db      3327  AAGTTTGGATTTGAATTCAGACAGAGAAATTCGCAATTAATCTTACCAATTATTA 3386
Qy      1093  ThrTyrAlaLeuSerSerValGlySerProLysAlaLysGluAlaLeuAsnMetLeuThr 1112
Db      3387  TCCATATGCCGTGTCCACAGTGGAGGCCCTTAAGACAGAGAGGCTTTGAATCTGTGATG 3446
Qy      1113  TrpArgAlaGluGlnGlyGlyMetGlnPheTrpValSerSerGluSerLysLeuSer 1132
Db      3447  CAGCATCATGAAGAAAGAGAGAGACACAGATCTGTGTGTATCATCAGACCTGCGCTCTCT 3506
Qy      1133  AspSerTrpGlnProArgSerLeuAspIleGluValAlaAlaTyrAlaLeuLeuSer 1152
Db      3507  GGGTCTCGACCGCGCTCCGTGACATCAAAAGCGTGTATGACCTGTGACCTGTGGCGAC 3566
Qy      1153  PheLeuGlnPheGlnThrSerGluGlyIleProIleMetArgTrpLeuSerArgGlnArg 1172
Db      3567  ACTCTGCAC-----CANGTGTGAGAGGAAATCCGGTTATGAGGTGGCTCATCAGCAAGAA 3623
Qy      1173  AsnSerLeuGlyGlyPheAlaSerThrGlnAspThrThrValAlaLeuLysAlaLeuSer 1192
Db      3624  AACACCTCGGAGAGGTTTGTATCACTAGAGATCTGTGTGTGGCCCTGAAGGCATTAATCT 3683
Qy      1193  GluPheAlaAlaLeuMetAsnThrGluArgThrAsnIleGlnValThrValThrGlyPro 1212
Db      3684  GAATTTTCAAGCTTGTGTGACAGAGAAATACAGATATCCAACTGACCGTCAAGGCGCC 3743
Qy      1213  SerSerPro----- 1215
Db      3744  GGCATCCCAAGATCATACACTTCGGAATGTAGACTCCAGAACTCTTTCTCTTCACAG 3803

```

```

Qy 1216 SerProLeuAlaValAlaGlnProMetAlaValAsnIleSerAlaAsnGlyPheGlyPhe 1235
Db 3804 GAAGAGCTTCAATGGCGTATGATCCATCACCCTGATATGTTTCTGGCAGATGGCTCGGAAATT 3863
Qy 1236 AlaIleCysGlnLeuAsnValAlaIleThrAsnValAlaSerGlySerSerArgArgArg 1255
Db 3864 GCTATTTGGCCAGCTTAATGATGACTATTAACCTGAAGGTTTCTTCTTAAAGACGG 3923
Qy 1256 ArgSerIleGlnAsnGlnGluAlaPheAspLeuAspValAlaValGluAsnLeuAsp 1275
Db 3924 AGATCTACCGAAATCAAGAAAGTTTTCATTTAAGCTCATTTG--AATTAATGAGGAC 3980
Qy 1276 AspLeuAsnIleValAspLeuAsnValCysThrSerPheSerGlyProGlyArgSerGly 1295
Db 3981 GACATTTAGTCACTCGAATCTGAATGTGCAAGATCCTGGGTTCAGAGAGACAGGC 4040
Qy 1296 MetAlaLeuMetGluValAsnLeuLeuSerGlyPheMetValProSerGluAlaIleSer 1315
Db 4041 ATGGCTCTTATGGAAGTGAACCTTCTCAGTGGCTTATGTCATCTTCAGATTCAATTCT 4100
Qy 1316 LeuSerGluThrValAlaValGluValAlaGluThrAspHisGlyLeuLeuAsnLeuLeuAsp 1335
Db 4101 CTGAGTGAAGACCTCGAAGAAAGTGAATATGATTAACGGAAATCTCAACCTTTATTTGAT 4160
Qy 1336 SerValAsnGluThrValPheCysValAsnIleProAlaValAlaAsnPheLeuValSer 1355
Db 4161 TCTGTAATGAATCCAGTATTTGTTGTTATATTTCCCATCTGAGAGACTACAAAGTTTCA 4220
Qy 1356 AsnThrGlnAspAlaSerValSerIleValAspTyrTyrGluProArgGlnAlaVal 1375
Db 4221 AATTTTCAGATGGTTCAGTATCTGTAATGACTACTATGAACCAAGAGACAGCGAGTG 4280
Qy 1376 ArgSerTyrAsnSerGluValAlaLeuSerSerCysAspLeuCysSerAspValGlnGly 1395
Db 4281 CGAAGCTACCAACACCGAGTGAAGCTGCTCCTCACTGATTAACCTCAGTCCAGAAC--AAC 4337
Qy 1396 CysArgProCysGluAspGlyAlaSerGlySerHisHisSerSerValIlePheIle 1415
Db 4338 TCGAAGTACACACAGGACGAGACGACGACCTCCCTGAGACTTCTTGAGGCTTCTGTC 4397
Qy 1416 PheCysPheLeuLeuLeuTyrPheMetGlu 1425
Db 4398 TTTTGCTCCGCTTCTGTTACTTGTGCAA 4427

```

RESULT 12  
 HSM805581 2938 bp mRNA linear PRI 12-JUL-2002

LOCUS Homo sapiens mRNA; cDNA DKFZp762L1111 (from clone DKFZp762L1111).  
 AL834478  
 ACCESSION AL834478.1 GI:21740252

VERSION  
 KEYWORDS  
 SOURCE Homo sapiens (human)

ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 1 (bases 1 to 2938)

REFERENCE Blum, H., Baurerachs, S., Mewes, H.W., Well, B. and Wiemann, S.  
 Direct Submission  
 Submitted (09-JUL-2002) 1, D-85764 Neuherberg, GERMANY  
 JOURNAL Clome from S. Wiemann, Molecular Genome Analysis, German Cancer  
 Research Center (DKFZ), Email: s.wiemann@dkfz-heidelberg.de;  
 sequenced by LMU (Ludwig Maximilians University,  
 Munich/Germany) within the cDNA sequencing consortium of the German  
 Genome Project.

FEATURES  
 TITLE This clone (DKFZp762L1111) is available at the RZPD in Berlin.  
 JOURNAL Please contact the RZPD: Ressourcenzentrum, Heuberweg 6, 14059  
 Berlin-Charlottenburg, GERMANY; Email: clone@rzd.de. Further  
 information about the clone and the sequencing project is available  
 at http://mips.gsf.de/proj/cDNA/.

FEATURES  
 SOURCE Location/Qualifiers  
 1..2938  
 /organism="Homo sapiens"

```

/mol_type="mRNA"
/db_xref="taxon:9606"
/map="15"
/clone="DKFZp762L1111"
/cisue_type="melanoma (Melo cell line)"
/clone_lib="762 (synonym: hmel12). Vector pSport1; host
DH10B; sites NotI + SalI"
/dev_stage="adult"
1..2938
/gene="DKFZp762L1111"
<1..2565
/feature="similarity to alpha-2-macroglobulin (Limulus sp.)"
/feature="similarity to alpha-2-macroglobulin (Limulus sp.)"
/codon_start=1
/product="hypothetical protein"
/protein_id="CAD39137.1"
/db_xref="GI:21740253"
/translations="VKSIVLMAASNDITMENVVELELYNTGYLGMPNRAVQV
CGLWVLDANLTDYIDGYDNEAVERMEVEHGTVIDHDSLSSPVRKHPET
MILDTNMGSRIVQEPVTVPSITSMVATGVEISDGLTTEVELOAQPFTF
LMLPVSVIRGEFALBITFNYLKDAEYVILEKDFDILMTSEINATGQTL
VPEDGATVLFIRPHEIPITVVALSPASDAITOMILVKAEGIEKSSGSLD
LTGNRSTLKTLSFSPRPNTVGSRRVOTITAGDVLGSLINGSLIAMPVCGSN
MIRPAPNITLIDLTCKKQITDNLKALSPKQYQRELYIREGESSAPENYPS
GSTWLSAFVLRCLFADPYIIDQNLHRTYMLKQHSNGBFMDPGVHSELQGG
NKPSPVLTAVITVSLGKRYOPNDIVQESIHLESEFSGISDNTLALITYALS
GSPKAKALNMLTAEQGMQFVSSKLSDSQPSRLDIEVAVALSHFLQFQ
TSRGIIDPHNRLLOTAELAVQPAVNIANGRPALCOLNVTYVNVKASGSRP
SPYKPIIDPHNRLLOTAELAVQPAVNIANGRPALCOLNVTYVNVKASGSRP
STONQAFIDVAVKENDLHNDLVNCTSFSGPERSGMLMEVNLGSPVPSAI
SLSEYKVEYDHGKLNLYDSVNETQFCVNIIPAVNFPVNTQDSVSIIDYEBR
QAVRSYNSKVKLSCDLCSDVOCRCEDGASGSHHSVIFPCFLFMELML"
2843..2848
/gene="DKFZp762L1111"
2859
polyA_site
BASE COUNT 887 a 529 c 620 g 902 t
ORIGIN
Alignment Scores:
Pred. No.: 5.54e-274 Length: 2938
Score: 4305.50 Matches: 834
Percent Similarity: 97.78% Conservative: 1
Best Local Similarity: 97.66% Mismatches: 2
Query Match: 58.59% Indels: 17
DB: Gaps: 1
US-10-020-095-4 (1-1428) x HSM805581 (1-2938)
Qy 592 ValAspLeuSerValAsnLeuMetAlaSerAsnAspIleThrMetGluAsnVal 611
Db 1 GTTGACAAAGTGTGAATCTGAATGCTCCTCAAGATTTCAATGCAAAATGGTGC 60
Qy 612 HisGluLeuGluLeuTyrAsnThrGlyTyrTyrLeuGlyMetPheMetCysSerPheAla 631
Db 61 CATTGAGTTGAACTTATTAACACAGATATATTATTTGCGCATGTTTCAATTTTGGCA 120
Qy 632 ValPheGlnGluCysGlyLeuTyrValIleThrAspAlaAsnLeuThrIleAspTyrIle 651
Db 121 GTTTTCAGAAAGTGAAGTGGATGATGACAGAACTTCAGAAAGATTTATTT 180
Qy 652 AspGlyValTyrAspAsnIleGluTyrAlaGluArgPheMetGluGluAsnGluGlyHis 671
Db 181 GATGCGTTTATGACAAATGCAAGATATGCTGAGAGGTTTATGAGAAATGAGAGCAT 240
Qy 672 IleValAspIleHisAspPheSerLeuGlySerSerProHisValArgIleHisPhePro 691
Db 241 ATTGTGATATTCATGATCTTTCTTGGGATGAGATCCACATGTCGAAAGCATTTTCCA 300
Qy 692 GluThrThrIleThrLeuAspThrAsnMetGlyTyrArgIleTyrGlnGluPheGluVal 711
Db 301 GAGACTTGATTTGGCTTACACCAACATGGGTTCCAGAGATTTTCAAGAAATTTGAAGTA 360

```



QY 712 ThrVal ProAsnSerIleThrSerThrProValAlaThrGlyPheValIleSerGluAspLeu 731  
 Db 361 ACTGTAACCTGATTCATCACTCTCTGGGGCTACAGGTTTGTGATCTCTGGAGCACTGG 420  
 QY 732 GlyLeuGlyLeuThrThrThrProValGluLeuGlnAlaPheGlnProPheIlePhe 751  
 Db 421 GGCCTTGGACTAACCACTACTCCAGTGGAGCTCCAAAGCCTTCCAAACATTTTTCATTTT 480  
 QY 752 LeuAsnLeuProTyrSerValIleArgGlyGluGluPheAlaLeuGluIleThrIlePhe 771  
 Db 481 TTGAATCTTCCCTACTCTGTATCAGAGGGAGAAATTTGCTTGGAAATTAATATATTC 540  
 QY 772 AsnTyrLeuLeuAspAlaThrGluValIleValIleIleGluLysSerAspLysPheAsp 791  
 Db 541 AATTAATTTGAAAGATCCACTGAGGTTAAGGTATCATGAGAAAGTACAAATTTGAT 600  
 QY 792 IleLeuMetThrSerSerGluIleAsnAlaThrGlyHisGlnGlnThrLeuLeuValPro 811  
 Db 601 ATTCTAATGACTTCAAGTAAATTAATTAATGCCAGGCCACAGCAGACCCCTTCTGGTTCCC 660  
 QY 812 SerGluAspGlyValaThrValLeuPheProIleArgProThrHisLeuGlyGluIlePro 831  
 Db 661 AGTGAAGATGGGGCACTGTTCTTTTCCATCAGGCCAACACATCTGGAGAGAAATTCCT 720  
 QY 832 IleThrValThrAlaLeuSerProThrAlaSerAspAlaValThrGlnMetIleLeuVal 851  
 Db 721 ATCACAGTACACAGCTCTTTCACCCAGCTTCTGATGCTATCACCCAGATGATTTAAGTA 780  
 QY 852 LysAlaGluGlyIleGluLysSerTyrSerGlnSerIleLeuLeuAspLeuThrAspAsn 871  
 Db 781 AAGGCTGAAGAAATTAAGAAATCATATCAATCAATCTTATTAACATGATGACAT 840  
 QY 872 ArgLeuGlnSerThrLeuLysThrLeuSerPheSerPheProProAsnThrValThrGly 891  
 Db 841 AGGCTACAGAGTACCTCGAAGAACTTTGAGTTCTCATTTCTCTTAATCAGAGTGGC 900  
 QY 892 SerGluArgValGlnIleThrAlaIleGlyAspValLeuGlyProSerIleAsnGlyLeu 911  
 Db 901 AGTGAAGAGTTACAGATCACTGCATATGAGAGATGTTCTTGTCCTTCATCAATGCTTA 960  
 QY 912 AlaSerLeuIleArgMetProTyrGlyGlyGlyGluGlnAsnMetIleAsnPheAlaPro 931  
 Db 961 GCCTCATGATTCGATGATGCTTATGCTGTGGTGAACAGAACTGATTAATTTTGTCTCA 1020  
 QY 932 AsnIleTyrIleLeuAspTyrLeuThrLysLysLysGlnLeuThrAspAsnLeuLysGlu 951  
 Db 1021 AATATTTCATTTTGGATTTATCTGACTAAAAGAAACACATGACACATTAATTTTGAAGA 1080  
 QY 952 LysAlaLeuSerPheMetArgGlnGlyTyrGlnArgGluLeuLeuTyrGlnArgGluAsp 971  
 Db 1081 AAGGCTCTTTCATTTATGAGGCAAGGTTACAGAGAACTTCTCATCAGAGGGAAGAT 1140  
 QY 972 GlySerPheSerAlaPheGlyAsnTyrAspProSerGlySerThrTyrLeuSerAlaPhe 991  
 Db 1141 GGCCTCTTTCAGGCTTTTGGGAATTAAGACCTTCTGAGGAGCACTTGTTGTACAGCTTT 1200  
 QY 992 ValLeuArgCysPheLeuGluAlaAspProTyrIleAspIleAspGlnAsnValLeuHis 1011  
 Db 1201 GTTTTAAGATGTTTCTGTAAGCCGATCTTACATGATTAATGATCAGAAATGTTTACAC 1260  
 QY 1012 ArgThrTyrThrTyrLeuLysGlyHisGlnLysSerAsnGlyGluPheTyrAspProGly 1031  
 Db 1261 AGAAACATACCTTGGCTTAAAGGACATCAGAAATCCACGGTGAATTTTGGGATCCAGGA 1320  
 QY 1032 ArgValIleHisSerGluLeuGlnGlyGlyAsnLysSerProValIleThrLeuThrAlaTyr 1051  
 Db 1321 AGGTATTCATCATGAGCTTCAAGGTGGCAATTAAGTCCAGTACACTTACAGGCTTAT 1380  
 QY 1052 IleValThrSerLeuLeuGlyTyrArgLysTyrGlnProAsnIleAspValGlnGluSer 1071  
 Db 1381 ATTGTAACTTCTCTCTCGGATTAAGAAAGTATCAGGCTTACATTAATGATGCAAGAGTCT 1440  
 QY 1072 IleHisPheLeuGluSerGluPheSerArgGlyIleSerAspAsnTyrThrLeuAlaLeu 1091

Db 1441 ATCCATTTTGGAGTCTGAATTCAGTAGAGAAATTTTACCAATTAATCTAGCCCTT 1500  
 QY 1092 IleThrTyrAlaLeuSerSerValGlySerProLysAlaLysGluAlaLeuAsnMetLeu 1111  
 Db 1501 ATAACTTATGCAATTTCTCATGCTGGGAGTCTCTTAAGCGAAAGAGCTTTGAATATGCTG 1560  
 QY 1112 ThrTyrArgAlaGluGlnGluGlyGlyMetGlnPheThrProValSerSerGluSerLysLeu 1131  
 Db 1561 ACTTGAGAGCGAACAAGAAAGGTGGCATGCAATCTGGGTGTATCAGAGTCCAACTT 1620  
 QY 1132 SerAspSerTyrGlnProArgSerLeuAspIleGluValAlaAlaTyrAlaLeuLeuSer 1151  
 Db 1621 TCTGATCTCTGGCAGCCACGCTCCCTGGATTTGAAATTTGAGCCTTATGCACTGCTCA 1680  
 QY 1152 HisPheLeuGlnPheGlnThrSerGluIleLeuProIleMetArgTyrLeuSerArgGln 1171  
 Db 1681 CACTTCTTACAAATTTTCAACATTTGAGGGAATCCCAATTAATGAGGTGGCTAAGCGGCA 1740  
 QY 1172 ArgAsnSerLeuGlyGlyPheAlaSerThrGlnAspThrThrValAlaLeuLysAlaLeu 1191  
 Db 1741 AGAAATAGCTTGGGTGGTGTTCATCTACATCTACAGATACCACTGGCTTTAAAGCTCTG 1800  
 QY 1192 SerGluPheAlaAlaLeuMetAsnThrGluArgThrAsnIleGlnValIleThrValThrGly 1211  
 Db 1801 TCTGAATTTGCAAGCCTAATGAAATACAGAAAGAACAAATATCAAGTACCGTGA 1860  
 QY 1212 ProSerSerProSerPro----- 1217  
 Db 1861 CCTAGCTCACCAAGTCCGTAAAGTTTCTGATGACACACACAAACCGCTTACTCTTCAG 1920  
 QY 1218 -----LeuAlaValAlaGlnProMetAlaValaAsnIleSerAlaAsnGlyPheGly 1234  
 Db 1921 ACAGAGAGCTTGGTGTGTACAGCCACGCAAGTAAATTTCCGCAAAATGTTTGA 1980  
 QY 1235 PheAlaIleCysGlnLeuAsnValValTyrAsnValLysAlaSerGlySerSerArgArg 1254  
 Db 1981 TTTGCTATTTGCACTCAATGTTGTATATGATGAAGGCTTGGGTCTTCTGAAGA 2040  
 QY 1255 ArgArgSerIleGlnAsnGlnGluAlaPheAspLeuAspValAlaValLysGluAsnLys 1274  
 Db 2041 CGAAGATCTATCCAAATCAAGAAAGCCTTGAATTAAGTGTGCTGTAAGAAATTA 2100  
 QY 1275 AspAspLeuAsnHisValaAspLeuAsnValCysThrSerPheSerGlyProGlyArgSer 1294  
 Db 2101 GATGATCTCAATCATGTGATTTGATGTGTACACCTTTTCCGGCCCGGTRGGAGT 2160  
 QY 1295 GlyMetAlaLeuMetGluValaAsnLeuLeuSerGlyPheMetValProSerGluAlaIle 1314  
 Db 2161 GGCATGGCTCTTATGGAAGTTAACCTATTAAGTGCCTTAAGTCCCTTGAAGCAATT 2220  
 QY 1315 SerLeuSerGluThrValLysLysValaGluTyrAspHisGlyLysLeuAsnLeuTyrLeu 1334  
 Db 2221 TCTCTGACCGAGACAGTAAAGAGTGAATGATCATGGAATAACTCAACTATTTTA 2280  
 QY 1335 AspSerValaAsnGluThrGlnPheCysValaAsnIleProAlaValaArgAsnPheLysVal 1354  
 Db 2281 GATTCGTGAAGAAACCAAGTTTGTATATATCTCTGCTGTGAAGAACTTTAAAGTT 2340  
 QY 1355 SerAsnThrGlnAspAlaSerValSerIleValAspTyrTyrGluProArgArgGlnAla 1374  
 Db 2341 TCAATTAACCAAGATGCTTCAAGTGTCCATATGATTAATTAAGCCAAAGGACAGGCG 2400  
 QY 1375 ValaArgSerTyrAsnSerGluValaLysLeuSerSerCysAspLeuCysSerAspValGln 1394  
 Db 2401 GTGAGAAATTAACAACTTGAAGTGAAGTGTCTCTCTGTGACCTTTGAGAGATGTCCAG 2460  
 QY 1395 GlyCysArgProCysGluAspGlyValaSerGlySerHisHisHisSerSerValIlePhe 1414  
 Db 2461 GGCCTCCGCTCTTGTGAAGATGAGACTTCAAGCTCCATCATCACTCTTCACTGATTTT 2520  
 QY 1415 IlePheCysPheLysLeuLeuTyrPheMetGluLeuThrLeu 1428

Db 2521 ATTTTCGTTCAAGCTTCTGACTTATGGAACCTTG3CTG 2562

RESULT 13  
LOCUS AK095888 2273 bp mRNA linear PRI 15-JUL-2002  
DEFINITION Homo sapiens cDNA FLJ38569 f18, clone HCHON2006459.  
ACCESSION AK095888  
VERSION AK095888.1 GI:21755236  
KEYWORDS oligo capping; f18 (full insert sequence).  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE  
AUTHORS Oshima,A., Takahashi-Fujii,A., Tanase,T., Imose,N., Takeuchi,K., Arita,M., Muraashino,K., Yuuki,H., Hara,H., Sugiyama,T., Irie,R., Otsubu,T., Sato,H., Makamatsu,A., Ishii,S., Yamamoto,J., Isono,Y., Kawai-Hio,Y., Saito,K., Nishikawa,T., Kimura,K., Yamashita,H., Matsuo,K., Nakamura,Y., Sekine,M., Kikuchi,H., Kanda,K., Magatsuuma,M., Murakawa,K., Kanehori,K., Sugiyama,A., Kawakami,B., Suzuki,Y., Sugano,S., Nagahari,K., Masuno,Y., Nagai,K. and Isegaki,T.  
TITLE NED0 human cDNA sequencing project  
JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 2273)  
AUTHORS Isegaki,T. and Yamamoto,J.  
TITLE Direct Submission  
JOURNAL Submitted (04-JUL-2002) Takao Isegaki, FLJ Project (HRI Team); 2-6-7 Kazusa-Kamatari, Kisarazu, Chiba 292-0812, Japan (E-mail:genomacenter.co.jp, tel:81-438-52-3975, Fax:81-438-52-3986) NED0 human cDNA sequencing project supported by Ministry of Economy, Trade and Industry of Japan; cDNA full insert sequencing: Research Association for Biotechnology (RAB); cDNA library construction: Helix Research Institute (HRI) (supported by Japan Key Technology Center etc.); 5'- & 3'-end one pass sequencing: RAB, HRI, and Biotechnology Center, National Institute of Technology and Evaluation; clone selection for full insert sequencing: HRI and RAB; annotation: HRI and RAB.  
location/Qualifiers  
1..2273  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="HCHON2006459"  
/cell\_type="Chondrocytes (HC)"  
/clone\_lib="HCHON2"  
/note="Cloning vector: pME18FLJ3-Primary culture, Chondrocytes"  
66..2063  
/note="unnamed protein product"  
/codon\_start=1  
/protein\_id="AAC04642.1"  
/db\_xref="GI:21755237"  
/translation="MQGPIILTAHLCTCTAALAVAGPRFLVTAEGIIIRPGNVTI  
GVEILHEHPSQVTVKAEILKTSANLTVSEVLEAGVFEKSGFKTLTSLPLNSDELY  
ELRYRTQDEILFNSRTSFKTRISVFIOTKALYKPEQVKEPILVILFSPFKY  
KTSNILLIKPKSNLLOQWLSQSDIGVISTFOLSHPIIGDMSIOVONDOVYOS  
FOVEYLVPKREYTLQTPLYCSMNSKHLNGITTKYTKYGRPKVDVLTLPPLSPMK  
KGNITKTFKINGSANFSFDEKMNWDSNGLESEYLDLSPGVEILTVTESVTGI  
SRNVSTVVFQKHDIILEFDYTVLKPSLNFATVTVTRADGQTLERRNNVIT  
VTORNATVEYSGNSGNDKMEAVOKINVTYPOSGTPIKEPFILEDSELMOLKAYFLS  
KSNAAVHSLFKPSKTYIOLKTRDENIKVSGPELVYSGNRLKELSYLVVSVOLVA  
VGNKSTMFSLTPENSWTPKACVIVYIEDGELISDYLKIPQVLVKNKIKLYMSKY  
KAERSEKSLISTVQPSDIYGVAVDKSVLMAASNDITMENVHLELELYNTGYLG  
MFMNSFAVFGCGJLWLTDAVLTYDVIQVYDNLFGTQEL"

BASE COUNT 712 a 431 c 462 g 668 t  
ORIGIN

Alignment Scores:  
Pred. No.: 5.9e-212 Length: 2273  
Score: 3360.00 Matches: 657  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0

Query Match: 45.73% Indels: 0  
DB: 9 Gaps: 0  
US-10-020-095-4 (1-1428) x AK095888 (1-2273)

Qy 1 MerglNglyProProleuleuThrrAlaAlaHileuleuCySValCyethrrAlaAlaLeu 20  
Db ATGACAGGCCACCGCTCCGACCGCCGCCACCTCTCTGCTGTGTACAGCGCGCTG 125  
Qy 21 AlavAlaProglyProAArgPheLeuValThrrAlaProglyIleIleAArgProgly 40  
Db GCGGTGCTCCGGGCTCCGCTCTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 185  
Qy 41 AsnValThrIleGlyValGluLeuGluHileCyAspSerGlnValThrrValLySAla 60  
Db AATGTGCTATTGGGGGTGGAGCTTGTGGAACATGCGCTTCAACAGGTGATGGAAGCG 245  
Qy 61 GluleuleuLyThrAlaSerAsnLeuThrrValSerValleuGluValGluGlyValPhe 80  
Db GAGCTGCTCAAGACAGCATCAACCTCACTGTCTGTCTGTGTGTGTGTGTGTGTGTGT 305  
Qy 81 GluLySerGlySerPheLyThrLeuThrrLeuProSerLeuProLeuAsnSerAlaAspGlu 100  
Db GAAAGAGCTCTTTTAAGACACTTACTTCTTCACTACTTCACTGACAGTGCAGATGAG 365  
Qy 101 IleTyGluLeuArgValThrrGlyArgThrrGlnAspGluIleLeuPheSerAsnSerThr 120  
Db ATTTATAGTACGT 425  
Qy 121 ArgLeuSerPheGluThrrLySArgIleSerValPheIleGlnThrrAspLySAlaLeuTyr 140  
Db CGCTTATCATTTAGACCAAGACAAATATCTGTCTTCACTTCACTTCACTTCACTTAC 485  
Qy 141 LySProLyGlnGluValLySArgIleValThrrLeuPheSerAspPheLySProTyr 160  
Db AAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCA 545  
Qy 161 LySThrSerLeuAsnIleLeuIleLySAspProLySerAsnLeuIleGlnGlnTrpLeu 180  
Db AAAAACTTTTAAACNTTCTCATTAAGAACCCCAATCAAAATTTGATCCAAAGTGTG 605  
Qy 181 SerGlnGlnSerAspLeuGlyValIleSerLyThrPheGlnLeuSerSerHileProIle 200  
Db TCACAACAAAGATGATCTGTGAGCATTTCCAAACCTTTTCAAGCATCTTCCATCAATA 665  
Qy 201 LeuGlyAspTrpSerIleGlnValGlnValAsnAspGlnThrrTyrrGlnSerPheGln 220  
Db CTTGTGACTGCTCATTCAGTTCAAGTGAATGACACACATCACTATCAATCAATTCAG 725  
Qy 221 ValSerGluTyrrValIleuProLySArgLeuValThrrLeuGlnThrrProLeuTyrrCySer 240  
Db GTTTCGAATATGATATTACCAAAATTTGAAGTACATTTGACAGACACATTAATTTGCT 785  
Qy 241 MetAsnSerLySAlaLeuAsnGlyThrrIleThrrAlaLySArgThrrTyrrGlySProVal 260  
Db ATGAATTTCTAAGCATTTAAATGATGATACATCAAGCAAGCAAGCAAGCAAGCAAG 845  
Qy 261 LyGlyLySAspValThrrLeuThrrPheLeuProLeuSerPheTrpGlyLySAlaIle 280  
Db AAGGAGACCTAACGCTTACATTTTACCTTTTGGGGAAGAAAGAAAGAAAGAAATTT 905  
Qy 281 ThrLyThrPheLySAlaGlnLySerAlaAsnPheSerPheAsnAspGluGluMetLyS 300  
Db ACAAAAACATTTAAGATTAATGATGATGCAAACTTCTCTTTTAATGATGAAGATGAAA 965  
Qy 301 AsnValMetAspSerSerAsnGlyLeuSerGluTyrrLeuAspLeuSerSerProglyPro 320  
Db AATGTATGATCTTCAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1025  
Qy 321 ValGluIleLeuThrrValThrrGluSerValThrrGlyLySArgArgAsnValSerThr 340  
Db GTAGAAATTTTAAACACAGTGAAGATGATGATGATGATGATGATGATGATGATGAT 1085



```

QY      341 AsnValPhePheylsGlnHisAspTyrIleIleGlnPhePheAspTyrThrValIleu 360
DB      1086 AATGTTCTTCAAGAAACATGATTCATCATGATTTTATTAATCTACTCTCTG 1145
QY      361 LysProSerLeuAsnPheThrAlaThrValIleValThrArgAlaAspGlyAsnGlnIleu 380
DB      1146 AAGCCATCTCTCACTTCAAGCCACTGTGAAAGTAACTGCTGATGGAACCAACAG 1205
QY      381 ThrLeuGlnGluArgAsnAsnValIleThrValThrGlnArgAsnTyrThrGln 400
DB      1206 ACTCTGAAAGAAAGAAATATGATGATCAATACAGTACACAGAGAACTTACTGAG 1265
QY      401 TyrTrpSerGlySerAsnSerGlyAsnGlnIleMetGlnAlaValGlnIleAsnTyr 420
DB      1266 TACTGAGCGGAGCTTAACAGTGAATAACAGAAATGAGAACTGTTCAGAAATTAATAT 1325
QY      421 ThrValProGlnSerGlyThrPheIleGlnPheProIleGlnIleAspSerGln 440
DB      1326 ACTGTCCCAAGAGAACTTTAAAGATTGAATTCCTCAATCTGAGAGATTCAGATGAG 1385
QY      441 LeuGlnLeuAlaTyrPheLeuGlySerIleSerSerMetAlaValHisSerLeuPhe 460
DB      1386 CTACAGTTGAAGCCCTATTCTTGGTAGTAAAGTACAGTGCAGTTCATAGTCTGTT 1445
QY      461 LysSerProSerIleThrTyrIleGlnLeuIleThrArgAspGlnAsnIleIleValGly 480
DB      1446 AAGTCTCTAGTAAAGCATCATCACTCAATCAAAAGATGAATAATTAAGGTGGGA 1505
QY      481 SerProPheGlnIleuValIleSerGlyAsnIleArgIleuIleGlnIleuSerTyrMetVal 500
DB      1506 TCCTCTTTGAGTGTGGTGTAGTGAACAAACGATGAAGAGTTAAGCTATTAAGTGA 1565
QY      501 ValSerArgGlyGlnIleuValIleValGlyIleGlnAsnSerThrMetPheSerLeuThr 520
DB      1566 GTATCCAGGGGACAGTTGGTGGCTGAGAGAAACAAATTCACAAATGTTCTTTAAAC 1625
QY      521 ProGlnAsnSerTrpThrProIleValIleValIleValTyrIleGlnIleuAspGly 540
DB      1626 CCAAGAAATCTTGAATCTCAAAAGCTGTGTAATGTGATTAATTAAGTGAATGATGG 1685
QY      541 GlnIleIleSerAspValIleuIleIleProValGlnIleuValPheIleAsnIleIleVal 560
DB      1686 GAATATTAAGTATGATTTCTTAATAATCTGTTCACTGTTTAAATTAAGATTAAG 1745
QY      561 LeuTyrTrpSerIleValIleValIleGlnProSerGlnIleuValSerIleuArgIleSerVal 580
DB      1746 CTAATATTGAGTAAGTGAAGCTGAACCATCTGAAGAAAGTCTCTTGAATCTCTG 1805
QY      581 ThrGlnProAspSerIleValGlyIleValIleValIleAspIleSerValIleuMetAsn 600
DB      1806 AACACAGCTGAGCTCCATAGTGGATTTGAGCTGTTGACAAAGTGAATCTGATGAT 1865
QY      601 AlaSerAsnAspIleThrMetGlnAsnValIleHisGlnIleuGlnIleuTyrAsnThrGly 620
DB      1866 GCCTCAATGAATTTCAATGAAATGAGTTCACAGAGTGAACCTTATTAACACAGGA 1925
QY      621 TyrTyrLeuGlyMetPheMetAsnSerPheAlaValPheGlnIleuGlyLeuTyrPyl 640
DB      1926 TATTAATTAGGCAATGTTCAATGAAATCTTTTGAAGTCTTTCAGAAATGAGATCTGGGA 1985
QY      641 LeuThrAspAlaAsnLeuThrIleAspTyrIleAspGlyValTyrAspAsn 657
DB      1986 TTGACAGATGCAAACTCAAGAGATTAATTAATGAGTGTATTAACAT 2036

```

```

RESULT 14
AX054953          4527 bp   DNA       linear   PAT 13-JAN-2001
LOCUS             AX054953
DEFINITION        Sequence 68 from Patent WO0073328.
ACCESSION         AX054953
VERSION           AX054953.1  GI:12228337
KEYWORDS          Caenorhabditis elegans
SOURCE            Caenorhabditis elegans
ORGANISM          Caenorhabditis elegans

```

```

REFERENCE
AUTHORS          Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditidae;
TITLE            Rhabditioidea; Rhabditidae; Peloderinae; Caenorhabditis.
JOURNAL          1
                Unc-5 constructs and screening methods
                Patent: WO 0073328-A 68 07-DEC-2000;
                Devgen NV (BE)

FEATURES
source           1..4527
                 /organism="Caenorhabditis elegans"
                 /mol_type="genomic DNA"
                 /db_xref="caxon:6239"

BASE COUNT       1313 a      948 c      1077 g      1189 t
ORIGIN
Alignment Scores:
Pred. No.:       2,276-112      Length:      4527
Score:           1854.50        Matches:      476
Percent Similarity: 51.23%      Conservative: 292
Best Local Similarity: 31.75%   Mismatches:   562
Query Match:     25.24%         Indels:       169
DB:              Gaps:         44

US-10-020-095-4 (1-1428) x AX054953 (1-4527)
QY      15 ValCysThrAlaAlaLeuAlaValAlaProGlyProArgPheLeuValThrAlaProGly 34
DB      82 GTGTCAACACCGCGCGCCGACGTTAAACACGCG---ACTTATATCTTGTGCGCCACACA 138
QY      35 IleIleArgProGlyGlyAsnValThrIleGlyValGlnIleuGlnIleCysProSer 54
DB      139 GTAAGTCAACCGGATCAACCGCTTCCTGTTGATGAACCTTTGAAACAGACTCGGAT 198
QY      55 GlnValThrValIleValAlaGlnIleuLeuThrIleAsnIleuThrValIleSerValIleu 74
DB      199 GAAGACATGATTTGCGATTTGAAGTGAACGAACTGAACGAAACGAGACGATTCGCTCGA 258
QY      75 GlnAlaGlnGlyValPheGlnIleuGlySerPheIleThrIleuProSerIleuPro 94
DB      259 GTTATTAAGCAATTTG---AAGCGGGAATCCGACAGACTGTGAGCTTAAGTGAATGCT 315
QY      95 LeuAsnSerAlaAsp-----GlnIleTyrGlnIleuValIleThrGlyArgThrGlnAsp 112
DB      316 GCTCAATCACTAATCTCTAAGGAAAGCTATTAAGCTCTACATCCCGGAGAGCTCTTAAC 375
QY      113 ---GlnIleLeuPheSerAsnSerThrArgIleuSerPheGlnIleuTyrArgIleSerVal 131
DB      376 GCAGAACTCATATTTGAACAAACGAAATGATTAAGTCAACAAAGCTTTATCAAGTT 435
QY      132 PheIleGlnThrAspIleAlaLeuTyrIlePheProIleGlnIleuValIlePheArgIleVal 151
DB      436 TTCAATCAACAGATGAGCAATCTATGCGCTGCTTCTTGTTCGATACCGGGCAATTT 495
QY      152 ThrLeuPheSerAspPheIleProTyrIleThrSerIleuAsnIleLeuIleLysAspPro 171
DB      496 GTAGTGAATGCGAATGTAACCTTAAGCTCGGAAACGCTACTATTAAATTTTGCATGCA 555
QY      172 LysSerAsnLeuIleGlnGlnIleThrIleuSerGlnIleuAspLeuGlyValIleSerIle 191
DB      556 AGCAGAAATCTGATTTCAACAGACGATGTGAATCTTGTGATGAGAGATTAACAGCGCG 615
QY      192 ThrPheGlnIleuSerSerHisProIleGlnIleuGlyAspTyrIleGlnIleuValGln----- 209
DB      616 GAGCTCCAACTTGCCGAAAGACCTCTCTGAGATGCTTTATTCAGAGTGAACCTCG 675
QY      210 -----ValAsnAspGlnThrTyrTyrGlnSerPheGlnValIleSerGlnTyrValIleuPro 227
DB      676 AATGAGTCCAGAGAAATCG-----TCATTACGATGATACATACATGCTGCTTCC 726
QY      228 LysPheGlnIleuValThrIleGlnIleProLeuTyrCysSerMetAsnSerIleHisLeuAsn 247
DB      727 AAATTTGAAGTGAATATCAAGACGCTCTGTTATTAATTAACGAT---GATTGTGCA 783

```

Oy	248	GLYThrIleThrAlaIleSerThrThrThrGlyLysProValLysGlyAspValThrIleuThr	267
Db	784	GTGTTGGTCGATGCAAAAGTACACATATATGAAAAAGCGTTGCTGGAAAGCCAAAGTATCT	843
Oy	268	Phe---LeuProLeuSerPheTrp-----	275
Db	844	CTAAGACTTCCATGGCACAGATGGCATGGATGTTCCCATATCATATTGACGAAATATGA	903
Oy	276	LysLysLysAsn-----IleThrLysThrPheLysIleAsn-----GlySer	289
Db	904	GTCAAGAAGAAAGAGAGACTGATGTGTGACGGCTACTGTAAAGCTGAACAGACAAAGAGAA	963
Oy	290	AlaAsnPheSerPheAsnAspGlnGluMetLysAsnValMetAspSerSerAsnGlyLeu	309
Db	964	GCGGCGAGTTGTTTCTTCGATATGATAGCTCAAGAGG-----	999
Oy	310	SerGluTrpLeuAspLysSerSerProGlyProValGluIleLeuThrThrValThrGlu	329
Db	1000	CACAAAGTTCCTACATGAAATGGGAGAGAGTTCATATTCGAATTTGTTGCTCCTCACTGAA	1055
Oy	330	SerValThrGlyIleSerArgAsnValSerThrAsnValPhePheLysGlnHisAspTrp	349
Db	1060	GACATCATGAAATCGAAGAAAGAAAGCGTACTCATCAATTTCAACGTTCCGAGAGAGGTC	1119
Oy	350	IleIleGluPhePheAspTrpThrThrValLeuLysProSerLeuAsnPheThrAlaThr	369
Db	1120	AAACTGTGATGTTGAAAMACAGGAGACACCTTCACCCAGCAGACTCATTAATATGTTGTA	1179
Oy	370	ValLysValThrArgAlaAspGlyAsn-----GlnLeuThrLeuGluLysValArgAsn	387
Db	1180	GTTGCACCTGAAACAAATGATGATGACATCCCGTGAAGGCACATTGGCCAAAGAGA----	1233
Oy	388	AsnValValIleThrValThrGlnArgAsnTrpThrGluTrpSerGlySerAsnSer	407
Db	1234	-----GTTCAAGTATCAACAATTCATCAACTATCA-----TACAATCATGACACTTCA	1281
Oy	408	GlyAsnGlnLysMetGlnAlaValGlnLysIleAsnTrpThrValProGlnSerGlyThr	427
Db	1282	TCTCTCAAGAGGAAAAAACAAMAAATTTGGAGAGTTGACGCTACAGGCACTCCGGTT	1341
Oy	428	PheLysIleGluPheProIleLeuGlnLysSerSerGluLeuGlnLeuLysAlaThrPhe	447
Db	1342	CTCACCTTCACCAACCAATCACTGACACTTCTGCTGCATTTGAGGCTCAT---TATGAT	1398
Oy	448	LeuGlySerLysSerSerMetAlaValHisSerLeuPheLys-----	461
Db	1399	ATTGGTGGCAAGACAACTTCACAGCCACCCCAATATATTCCTGCTATATGTGAGGCA	1458
Oy	462	-----SerProSerLysThrTrpIleGlnLeuLysThrArgAspGluAsn---IleLys	478
Db	1459	GCCGTGACCCCAACAGAGATTTCTCTGACGTTTGGCTGTATTAAGAGAGAGCTTTGAT	1518
Oy	479	ValGlySerProPheGlnLeuValValSerGlyAsnLysAsnGlyLeuGlnLeuSerTrp	498
Db	1519	GTAGGAAAGAGCTTAGCTTCTCGTGAAGGCTACCGCGCTGTCCAAATTACATTCAT	1578
Oy	499	MetValValSerArgGlyGlnLeuVal-----AlaValGlyLeuGlnAsn	513
Db	1579	CAGGTCACTCTCTGTCGAACATTTGTGTCTCAACAAATGACTGTTATTTCCGAACAT	1638
Oy	514	SerThrMetPheSerLeuThrProGluAsnSerTrpThrProLysAlaCysValIleVal	533
Db	1639	GCGACCAATT---AGCTTCCCAAGCAACGGCTAAACATGAGCCCAAGTACAGGTGATGCTC	1695
Oy	534	Tyr---TrpIleGluAspAspGlyGlnIleIleSerAspValLeuLysIleProValGln	552
Db	1696	TATGCCATTATATGAGAGTATGACAGAAAGATTCTGTGTATGACACTTGACTTTAAAGTAAA	1755
Oy	553	LeuValPheLysAsnLysIleLysLeuLysTrpSerLysValLysAlaGluProSerGlu	572
Db	1756	GGAATATTCCAAACAAAGTGGCACTATCAATTGCAACAAACAGCACTGAGCGACGAGACAG	1815
Oy	573	LysValSerLeuArgLysSerValThrGlnProAspSerIleValGlyIleValAlaVal	592

Db 1816 AATGTGAAGTTTAAAGCTCA---TCGACAGAACTCTTTGTGGGCTTCCTTGCGTA 1872  
 Oy AApYsSerValaAenLeuMetAsnAlaSerAsnApriLeThiMetGluAenValaHis 612  
 Db 1873 GATCAGAGTGTTCTCTTTGTAATAAAACGGCAGATGATATCACTAGAGAAAAAGTTGACAA 1932  
 Oy 613 GluLeuGluLeuTyTzAnThrGlyTyTzTyLeuGlyMetHe----- 626  
 Db 1933 GACTTGGAGATTATGTACTTAATATGTTGGAGAGGATTCGGGGGTCCACGGCCATGG 1992  
 Oy 627 -----Met 627  
 Db 1993 GAGCCAAATTGATAGAAAGAAAGSTCCATTTCGAGACCTTCGTGGGGAATTGGAGGAGT 2052  
 Oy AAsnSerPhealAvalPheGlnGluCyGlyLeuTrpValLeuThrAspAlaAenLeuThr 647  
 Db 2053 GATGCTCAATCAATCTTTTCTTAATGCGTGGCTGTAGATTTTGGACCGACGCGCTCTTGTC 2112  
 Oy LysAsp-----TyTleAspGlyValTyTzAspAsnAla 658  
 Db 2113 CGCGACGCTCAACGTGAATTTATGTCTAGTAATGATATGATGGATGTCTCCAGGAATGGT 2172  
 Oy GluTyTzAlaGluArgPheMetGluGluAenGluGlyHisIleValAspIleHisAspPhe 678  
 Db 2173 GAACCCGA-----TTCGCCGCTCCGCCAATGGC----- 2202  
 Oy SerLeuGlySerSer-----ProHisValArgLysHisPheProGluTrpIle 695  
 Db 2203 -----GGTGCTGCTCCCTCCACCGCCGACAGTTTGAAAGTTCTTCCACACACCTGGATT 2256  
 Oy 696 TrpLeuAspThrAsnMetGlyTyTzArgIleTyTzGlnGluPheGluValThrValProAsp 715  
 Db 2257 TGGTCGAGATCCAAAC---TCTACTAGTGGAGAAAGTTAGATGGAATAATGGAAGCTCCGCGAC 2313  
 Oy 716 SerIlePheSerTrpValAlaThrGlyPheValIleSerGluAspLeuGlyLeuGlyLeu 735  
 Db 2314 ACCATCACTCTGTGGTGGCTCTCCACTTTTGCAATCATAGAGAAATGGCTCGCGGTG 2373  
 Oy 736 ThrThrThrProValGluGlnAlaPheGlnProPhePheIlePheLeuAenLeuPro 755  
 Db 2374 GCTCCACACGACTTCAAAACTTCGCGGTCTTCGTCATTTTATTCACACTCATCTTCCA 2433  
 Oy 756 TySerValIleArgGlyGluGluPheAlaLeuGluIleThrIlePheAsnTyLeuLys 775  
 Db 2434 TATCTGTCCGCTGGAGAAAAGTTTGCACTTTGGTGTCTGTTTTCATTTATATGCA 2493  
 Oy 776 AspAlaThrGluValLysValIleIle-----GluLysSerAspLysPheAspIleLeu 793  
 Db 2494 AAGGAGCAAGTGTACAGTGACTGTAAGATGATATMAACCTGTGGTTACGATCTCTG 2553  
 Oy 794 MetThrSerSerGluIle-----AsnAlaThrGlyHisGlnGlnThrLeuLeuVal 810  
 Db 2554 AAGAAAGATGGAACCTAGTAGTAAGCCGCGACGAGGTTGCCAACAATACTTCGATATGTC 2613  
 Oy 811 ProSerGluAspGlyAlaThr-----ValLeuPheProIleArgProThrHisLeu 827  
 Db 2614 TCCGTGTGGTGGTAGGACATCGAAGACAGTGTACTTCCCATATGTGCGGTCTTCAT 2673  
 Oy 828 GlyGluIleProIleThrValThrAlaLeuSerProThrAlaSerAspAlaValThrGln 847  
 Db 2674 GGAAGAGATCCCGGTACATATAAAGCGTATTCGATCCCAAGTGGTATGACATAGAGATG 2733  
 Oy 848 MetIleLeuValLysAlaGluGlyIleGluLysSerTySerGlnSerIleLeuLeuAsp 867  
 Db 2734 AACTTGAAGTGTATCCACAAGGATATATTAAGTATGATAGAAATATTCATTTGTCATTGAC 2793  
 Oy 868 LeuThrAspAsnArgLeuGlnSerThrLeuLysThrLeuSerPheSerPheProAsn 887  
 Db 2794 CTCAACATATACCTTCGGAATTCCTC---AAAATCTTAGGCTTATCTGGCCAAACGAC 2850  
 Oy 888 ThrValThrGlySerGluArgValGlnIleThrAlaIleGlyAspValLeuGlyProSer 907





QY 685 HisValArgLysHisPheProGluThrTrpIleThrLeuAspThrAsnMetGlyTyrArg 704  
 DB 2257 ACAATTCGAAAGTCTTCCACACACTGGATTGGCTCGAATCTCAAC---TCTACTACT 2313  
 QY 705 IleTyrGlnGluPheGluValThrValProAspSerIleThrSerTrpValAlaThrGly 724  
 DB 2314 GGAGAAAGTTGAGATGAAATTGAAAGTCCGACACCATCATCTCGTGGTCCGCTTCCACT 2373  
 QY 725 PheValIleSerGluAspLeuGlyLeuGlyLeuThrThrProValGluLeuGlnAla 744  
 DB 2374 TTTCGAATCATAGAGAAATAGTCTCGGTGTGGCTCCAAAGCACTCAAAACTCGCGCG 2433  
 QY 745 PheGlnProPhePheIlePheLeuAsnLeuProTyrSerValIleArgGlyGluLeuPhe 764  
 DB 2434 TTTCCTCAATTTTATTCATCATCTTCATATGCTGCTCGTCCGTGAGAGAAAGTTT 2493  
 QY 765 AlaLeuGluIleThrIlePheAsnTyrLeuLysAspAlaThrGluValValIleIle 784  
 DB 2494 GCACCTTTGGTCTGCTTTTCAATTTATGAAAGAGCAAGATGTAACAGTACTG 2553  
 QY 785 -----GluLysSerAspLysPheAspIleLeuMetThrSerSerGluIle----- 799  
 DB 2554 AAGTATGATAAAGACTCTGTGTTACATCTCTGAGAGAAAGATGAACTGTATAGGCGT 2613  
 QY 800 AsnAlaThrGlyHisGlnGlnThrLeuValProSerGluAspGlyAlaThr----- 817  
 DB 2614 GACGAGGTGGCCCAAAATGCTGATTTGCTCGCTGCTGAGAGAAATGCAAA 2673  
 QY 818 ---ValLeuPheProIleArgProThrHisLeuGlyGluIleProIleThrValThrAla 836  
 DB 2674 GCAGGTACTTCCCAATGTGCGCTTCATTTGAGAGATCCCGTCAATTAAGCGT 2733  
 QY 837 LeuSerProThrAlaSerAspAlaValThrGlnMetIleLeuValLysAlaGlyIle 856  
 DB 2734 ATTCGATCCCAAGGTGGTATGACAGAGATGAAGTCAAGTTCATCAACAGATAT 2793  
 QY 857 GluLysSerTyrSerGlnSerIleLeuAspLeuThrAspAsnArgLeuGlnSerThr 876  
 DB 2794 AAGAGATAGATAATATTCATTTGCTGACTTCACTCAACATTAATCTTCGAGATTTCCT 2853  
 QY 877 LeuLysThrLeuSerPheSerPheProProAsnThrValThrGlySerGluArgValGln 896  
 DB 2854 ---AAAATCTTGACCTTATCGGCCCAAGACGTGGTGTGATGATCCGAGAGCAAA 2910  
 QY 897 IleThrAlaIleGlyAspValLeuGlyProSerIleAsnGlyLeuAlaSerIleArg 916  
 DB 2911 TTAGATGATTTGGTATGATGATGAGACCAAGTCTTAACATGCTCATAGCTGTACAA 2970  
 QY 917 MetProTyrGlyCysGlyGluGlnAsnMetIleAsnPheAlaProAsnIleTyrIleLeu 936  
 DB 2971 ATGCCATATGAGATGAGAGACAGATATGTTGAATCTGTTCCAAATATTTTGGTGTCT 3030  
 QY 937 AspTyrLeuThrLysLysLeuLeuThrAspAsnLeuLysGlyLysAlaLeuSerPhe 956  
 DB 3031 AAGTCTTAAGAGCAAAATATGAAATGAGACCAATTAGAGCAAAAGCTCTCAAGTTT 3090  
 QY 957 MetArgGlnGlyTyrGlnArgGluLeuLeuTyrGlnArgGluAspGlySerPheSerAla 976  
 DB 3091 ATTCAGCAAGGAAATTCAGAGAGTTGACGTACAAACGCTGACCAACTCTTTCAGCT 3150  
 QY 977 PheGlyAsnTyrAspProSerGlySerThrTyrLeuSerAlaPheValLeuArgCysPhe 996  
 DB 3151 TTTCGAGATAGTGAATGAAGCTGTTCACAGTGGCTACCGCATTTGTGCTGATCGTTC 3210  
 QY 997 LeuGluAlaAspProTyrIleAspIleAspGlnAsnValLeuHisArgThrTyrThrTrp 1016  
 DB 3211 CATCAAGCCCAAAATATCCATTCGATCTCAATGTTATTTCCGAGACAGTCCCATTC 3270  
 QY 1017 LeuLysGlyHisGlnLysSerAsnGlyGluPheTrpAspProGlyArgValIleHisSer 1036  
 DB 3271 TTGAATCTCAAAATGAGATGCTGTGATTCGCGCAAGCTGTGATTAATCATCAAG 3330  
 QY 1037 GluLeuGlnGlyGly---AsnLysSerProValThrLeuThrAlaTyrIleValThrSer 1055

DB 3331 GATATGAGGAGGAGGTCTCAAGACCGGGGTGTAGCTCTCAAGCGCAATTTGTGTATTCG 3390  
 QY 1056 LeuLeuGlyTyrArgLysTyrGlnProAsnIleAspValGlnGluSerIleHisPheLeu 1075  
 DB 3391 ATTTTG-----GAGATGAGATGAGAAATGAGAAAGCGGTGACGTATTTG 3435  
 QY 1076 GluSerGluPheSerArgGlyIleSerAspAsnTyrThrLeuAlaLeuIleThrTyrAla 1095  
 DB 3436 GAAAGCAATTTGATGAGATGATCTGGAATAGCGTATACAAAGGCTGTAGTGGCATATGCA 3495  
 QY 1096 LeuSerSerValGlySerProLysAlaLysGluAlaLeuAsnMetLeuThrTrpArg--- 1114  
 DB 3496 CTTCAATTGGCAAAAGATGACCAAGCTGAAAGAGCATTCGAGAAATTCAGAAACATTAAG 3555  
 QY 1115 ---AlaGlnGlnGlyGlyMetGlnPheTrpValSerSerGlu-----SerLysLeu 1131  
 DB 3556 ATCGTTGAGAAAGTGGAGATGTGAAGTTTGCCACTGCTCAAGAAAGATGAGAAACTA 3615  
 QY 1132 SerAspSer-----TrpGlnProArgSerLeuAspIleGluValAlaAlaTyr 1147  
 DB 3616 AAGAAATGAGAGCGTATATGATTCAGGCCGCTCCAGTATGACATTGAACACTACCTTAC 3675  
 QY 1148 AlaLeuLeuSerHisPheLeuGlnPheGlnThrSerGluGlyIleProIleMetArgTrp 1167  
 DB 3676 GCAGTCTTTTCTTACCTGCGCCCAAAATCAAACTGCAAGTCTCTCAATTAATTCGTTGG 3735  
 QY 1168 LeuSerArgGlnArgAsnSerLeuGlyLysPheAlaSerThrGlnAspThrThrValAla 1187  
 DB 3736 TTAGTATCCCAACGCAAGACTAGAGAGATTCCTTCACCTCAAGACAGCTGTGCGCT 3795  
 QY 1188 LeuLysAlaLeuSerGluPheAlaLeuMetAsnThrGlnArgThrAsnIleGlnVal 1207  
 DB 3796 CTTCAAGTCTCTCTCTTACGACGACGTCACTTATTCGACAAACACAAAGTAAAGTA 3855  
 QY 1208 ThrVal---ThrGlyProSerSerProSer-----ProLeuAlaVal 1220  
 DB 3856 ACAATTCACAGGAAACATATCATCATTTGATATCAATAGAAAGCAATTTGTT 3915  
 QY 1221 ValGlnProMet-----AlaValAsnIleSerIleAsnGlyPhe 1233  
 DB 3916 CTCCAATCTTATCAACTATCTTCATGATGATGATGAGATTCATTAATGACCAATGAACT 3975  
 QY 1234 GlyPheAlaIleCysGlnLeuAsnValValTyrAsnValLysAlaSerGlySerSerArg 1253  
 DB 3976 GGTGTGCTTCGCTCACTTGAATTTATCTTAC----- 4008  
 QY 1254 ArgArgArgSerIleGlnAsnGlnGluAlaPheAspLeuAspValAlaValLysGluAsn 1273  
 DB 4009 TACAGGACTCTTTGAATGAGAGATGCTCGTTCCTTCTGCTCCAGAGATCAAGAAATTA 4068  
 QY 1274 LysAspAspLeuAsnHisValAspLeuAsnValCysThrSerPheSerGlyProGlyArg 1293  
 DB 4069 AGAGCTGGA---AATCGCACTCAATTTGATTTGTTGCCAATATATACCCGCTTGGAAAG 4125  
 QY 1294 SerGlyMetAlaLeuMetGluValAsnLeuSerGlyPheMetValProSerGluAla 1313  
 DB 4126 TCTAATAGGCTCTTAGCCGAATATGATGCTTAGGTGTACCCGATTTGATGCGGAAAG 4185  
 QY 1314 IleSerLeuSerGluThrValLys-----LysValGluTyrAspHisGly 1328  
 DB 4186 GTGCACTACTTGAACAAGATTTGAGAGATTTGCAAAAGGTGAGAGTGAAGAAAGCAACT 4245  
 QY 1329 LysLeuAsnLeuTyrLeuAspSerValAsnGluThrGlnPheCysValAsnIleProAla 1348  
 DB 4246 AAGAGCAAGTATATTTCAACCGCTCGGTGCGCGTCCGATATGCTCTCGTTAATACTCT 4305  
 QY 1349 ValArgAsnPheLysValSerAsnThrGlnAspAlaSerValSerIleValAspTyrTrp 1368  
 DB 4306 GATGCACTTATCAAGTTGCGATCAAAACCTGCAACCTTCCGATTAAGTATTAATAC 4365  
 QY 1369 GluProArgGlnAlaValArgSerTyrAsnSerGluValLysLeuSerSerCysAsp 1388



Db 4366 GATCCGAGAGCACTCAAAATGACATACGCCGCCAAGCAACACGATCGCTGCAGAG 4425  
 QY 1389 LeuCyserAspValGlnGlyCysArgPro 1398  
 Db 4426 AATGTGGG-----GAAGACTGTTGGCCA 4449

Search completed: January 17, 2004, 08:32:16  
 Job time : 1145 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: January 16, 2004, 20:57:28 ; Search time 753 Seconds

(without alignments)  
5119.253 Million cell updates/sec

Title: US-10-020-095-4

Perfect score: 7348

Sequence: 1 MGGPPLTAAHLVCVTAA.....HSSVIFFCFKLYFMEIWL 1428

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 2552756 segs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seg length: 0  
Maximum DB seg length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DBV=xld  
-Q=/cgn2.1/USPro\_spool\_p/US10020095/rnuc.15012004.163821.18701/app.query.fasta.1.1607  
-DB=N.Geneseq.150un3 -Qfmt=fastag -SUFFIX=mg -MINMATCH=0.1 -LOOPEL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=biosum62 -TRANS=human40.cdi  
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=pco -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000  
-USER=US10020095 @CCN 1.1 603 @runat.15012004.163821.18701 -NCPU=6 -ICPU=3  
-NO\_MAP -LARGEOUTERY -NEG\_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREDS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FCAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

N.Geneseq.19Jun03:\*  
1: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT:\*  
2: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT:\*  
3: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT:\*  
4: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT:\*  
5: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT:\*  
6: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT:\*  
7: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT:\*  
8: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT:\*  
9: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT:\*  
10: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:\*  
11: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT:\*  
12: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT:\*  
13: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT:\*  
14: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:\*  
15: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT:\*  
16: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT:\*  
17: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT:\*  
18: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT:\*  
19: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:\*  
20: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:\*  
21: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:\*  
22: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:\*  
23: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:\*  
24: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:\*  
25: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2003.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed,

and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	7331	99.8	4369	25	AAD49434	Human r150 gene #1
2	7326.5	99.7	4335	24	ABL49815	Human platelet all
3	7326.5	99.7	4761	24	ABO79864	Human CD109 K1 pro
4	7326.5	99.7	5882	25	AAD49440	Human blood cell s
5	7326.5	99.7	5895	24	ABO79966	Human CD109 K1-H7
6	7318.5	99.6	5883	25	AAD49435	Human r150 DNA #1.
7	7317.5	99.6	4335	24	ABL49816	Human platelet all
8	7317.5	99.6	4761	24	ABO79965	Human CD109 K1 var
9	7317.5	99.6	5895	24	ABO79967	Human CD109 K1-H7
10	7307.5	99.4	4335	21	AAA60199	Hydrophobic domain
11	7307.5	99.4	4473	21	AAA62010	Hydrophobic domain
12	7081	96.4	4146	25	AAD49436	Human blood cell s
13	7068.5	96.2	4197	25	AAD49437	Human r150 DNA #2.
14	5470	74.3	3535	24	ABO79968	Human CD109 K15 pr
15	5461	74.3	3535	24	ABO79969	Human CD109 K15 va
16	1854.5	25.2	4527	22	AAO90951	C. elegans alpha-2
17	1850	25.2	4560	22	AAO90952	C. elegans alpha-2
18	1819.5	24.8	1448	24	ABL90478	Human polynucleoti
19	1767.5	24.1	1459	22	ABA08825	Human secreted pro
20	1533	20.9	5816	21	AAO74484	Human ORFX ORF3 p
21	1461.5	19.9	2403	24	ABV77330	Alpha-1, proteina
22	1441	19.6	4577	21	AAA52113	Human alpha-2-macr
23	1440	19.6	4425	22	AAO90944	Human alpha-2-macr
24	1440	19.6	4577	22	AAD07359	Human DNA encoding
25	1440	19.6	4577	24	AAD45346	Human anti-microbi
26	1440	19.6	4577	24	ABN95738	Gene #2236 used to
27	1440	19.6	4577	24	ABK64513	Human benign prost
28	1440	19.6	4577	24	ABK12952	DNA encoding human
29	1440	19.6	4577	24	ABK24095	Human alpha2 macro
30	1440	19.6	4578	25	ABX74411	Human cDNA sequenc
31	1440	19.6	4578	25	ABX74412	Human cDNA sequenc
32	1440	19.6	4809	24	AAO94923	Human DNA sequence
33	1438	19.6	4527	22	AAH74529	A human alpha-2 ma
34	1438	19.6	4527	22	AAH74530	A human alpha-2 ma
35	1430	18.5	3612	23	ABL19705	Drosophila melanog
36	1428.5	19.4	4569	12	AAO11176	Clone encoding rec
37	1425	19.4	4599	12	AAO11581	Sequence encoding
38	1422	19.4	4488	24	ABO59327	Human alpha-2-macr
39	1412.5	19.2	4595	24	ABK63624	Rat sequence diffe
40	1406	19.1	4278	23	ABL30183	Drosophila melanog
41	1401	19.1	4487	24	ABK92035	DNA encoding novel
42	1400	19.1	4615	23	ABN92193	DNA encoding novel
43	1400	19.1	4615	24	ABN97232	Gene #3730 used to
44	1376.5	18.7	4660	24	ABO93897	Human pregnancy zo
45	1376.5	18.7	6324	23	ABL19704	Drosophila melanog

#### ALIGNMENTS

RESULT 1  
AAD49434 standard; DNA; 4369 BP.  
XX AAD49434;  
AC  
XX  
DT 24-MAR-2003 (first entry)  
XX  
XX Human r150 gene #1.  
DE  
XX  
XX Human r150 gene #1.  
KW Human; GPI-anchored TGF-beta1 binding protein; r150 protein; cancer;  
KW glycosylphosphatidyl inositol; transforming growth factor-beta1;  
KW therapy; gene; ds.  
XX  
XX Homo sapiens.  
OS  
XX  
XX  
FH Key Location/Qualifiers

FT CDS 26..4312  
 FT /tag= a  
 FT /product= "Human r1520 protein #1"  
 FT /transl\_except= (pos:2132..2134, aa:Xaa)  
 FT /note= "Xaa corresponds to Ser, Tyr"  
 XX  
 XX WO200285942-A2.  
 XX  
 XX 31-OCT-2002.  
 XX  
 XX 24-APR-2002; 2002WO-CA00560.  
 XX  
 XX 24-APR-2001; 2001US-285713P.  
 XX 14-FEB-2002; 2002US-356163P.  
 XX  
 XX (UYMC-) UNIV MCGILL.  
 XX  
 XX Philip A, Tam B;  
 XX WPI; 2003-093100/08.  
 XX P-PSDB; AAE32012.  
 XX  
 XX Novel transforming growth factor (TGF)-beta 1 binding reagent which  
 XX comprises r150 protein which acts as accessory receptor of TGF-beta,  
 XX useful for negatively modulating TGF-beta activity, and thus for  
 XX treating cancer -  
 XX  
 XX Claim 7; Page 91-93; 127pp; English.  
 XX  
 XX The invention relates to novel transforming growth factor (TGF)-beta1  
 XX binding reagent which comprises a glycosylphosphatidy1 inositol (GPI)-  
 XX anchored TGF-beta1 binding protein referred to as r150 which acts as  
 XX an accessory receptor of TGF-beta. The invention is used for negatively  
 XX modulating TGF-beta activity, and thus for treating conditions  
 XX characterised by overproduction of TGF-beta, such as cancer. Antisense  
 XX molecules of the invention are used for increasing TGF-beta availability  
 XX and increase graft success. The present sequence is a gene encoding  
 XX human r150 protein.  
 XX  
 XX Sequence 4369 BP; 1301 A; 842 C; 932 G; 1293 T; 1 other;  
 SO  
 Alignment Scores:  
 Pred. No.: 0 Length: 4369  
 Score: 7331.00 Matches: 1425  
 Percent Similarity: 99.86% Conservative: 1  
 Best Local Similarity: 99.79% Mismatches: 2  
 Query Match: 99.77% Indels: 0  
 DB: 25 Gaps: 0  
 US-10-020-095-4 (1-1428) x AAD49434 (1-4369)  
 QY 1 MetGInGlyProProLeuLeuThrAlaAlaHisLeuLeuCyValCyThrAlaLeu 20  
 DB 26 ATGCAAGGCGCCACGCTCTGACCGCGCCACCTCTCTGCTGACCGCGCGCTG 85  
 QY 21 AlaValAlaProGlyProArgPheLeuValThrAlaProGlyIleIleArgProGly 40  
 DB 86 GCCGTGCTCCCGGCGCTCGGTTCTGTGACAGCCCGAGGATCATCAGGCCCGGAGA 145  
 QY 41 AsnValThrIleGlyValGluLeuLeuGluHisCysProSerGlnValThrValIleAla 60  
 DB 146 AATGTGACTATTGGGGTGGAGCTTCTGGAAACATGCGCTTCACAGGTGACTGAAGCG 205  
 QY 61 GluLeuLeuLeuThrAlaSerAsnLeuThrValSerValLeuGluAlaGluGlyValPhe 80  
 DB 206 GAGTGTCTCAAGACAGCATCAAACTCATCTGTCTGTCTGCGAAGCAGAGAGTCTTT 265  
 QY 81 GluLeuGlySerPheLeuThrLeuThrLeuProSerLeuProLeuAsnSerAlaAspGlu 100  
 DB 266 GAAAGAGGCTTTTAAAGACCTTACTCTTCCATCATCACTCTGAAACAGTGCAGATGAG 325  
 QY 101 IleTyrGluLeuArgValThrGlyArgThrGlnAspGluIleLeuPheSerAsnSerThr 120

DB 326 ATTATGAGCTACGTTGTAACCGGACGTACCCAGATGAGATTTTATTTCTTAATGATACC 385  
 QY 121 ArgLeuSerPheGluThrLysArgIleSerValPheIleGlnThrAspLysAlaLeuTyr 140  
 DB 386 CGCTTATCATTTTGAACCAAGAAATATCTGTCTTCAATCAAAACAGACAGGCTTATAC 445  
 QY 141 LysProLysGlnValLysPheArgIleValThrLeuPheSerAspPheLysProTyr 160  
 DB 446 AACCCAAAGCAAGAGGAAGTTCCGATGTGATCACTCTTCAGATTTTAAAGCCCTTAC 505  
 QY 161 LysThrSerLeuLeuIleLeuIleLysAspProLysSerLeuLeuIleGlnIleProLeu 180  
 DB 506 AAAACCTTTTAAACATTCATTAAGACCCCAAAATCAATTTGATCCAAACGTGGTTG 565  
 QY 181 SerGlnIleSerAspLeuGlyValIleSerLysThrPheGlnLeuSerSerHisProIle 200  
 DB 566 TCACACAAAGTATCTTGAGATCATTTCCAAAATCTTTGACGTATCTTCCATCCATTA 625  
 QY 201 LeuGlyAspTyrPheSerIleGlnValGlnValAsnAspGlnThrTyrTyrGlnSerPheGln 220  
 DB 626 CTGTGTACTGTGCTATTCATTAAGTTCAAGTGAATGACACAGACATATCATCATTTCCAG 685  
 QY 221 ValSerGluTyrValLeuProLysPheGluValThrLeuGlnThrProLeuTyrCysSer 240  
 DB 686 GTTTCAGAAATGTATTAACCAAAATTTGAGTACTTTGACACCATTTATATGTCTT 745  
 QY 241 MetAsnSerLysHisLeuAsnGlyThrIleThrAlaLysTyrThrTyrGlyLysProVal 260  
 DB 746 ATGAAATTTCAAGATTTTAAATGTATCATCAACGCAAAAGTATCATATGGAAAGCCAGTG 805  
 QY 261 LysGlyAspValThrLeuThrPheLeuProLeuSerPheTyrGlyLysLysAsnIle 280  
 DB 806 AAAGAGACGTAAACGCTTACATTTTAACTTTATCTTTGAGGAAAGAAAGAAATATTT 865  
 QY 281 ThrLysThrPheLysIleAsnGlySerAlaAsnPheSerPheAsnAspGluGluMetLys 300  
 DB 866 ACAAAAACATTTAAGATTAATGATGATCGCAACCTTCTTTAATATGATGAAGAGATGAAA 925  
 QY 301 AsnValMetAspSerSerAsnGlyLysSerGluTyrLeuAspLeuSerProGlyPro 320  
 DB 926 AATGTATATGATTTCTTCAATGATGATCTTTCGAATACCTGATATCTTCCCTGAGACA 985  
 QY 321 ValGluIleLeuThrThrValThrGluSerValThrGlyIleSerArgAsnValSerThr 340  
 DB 986 GTGAATTTTAAACCAAGTACAGATCATGTTACAGTATTTCAAGTAATATGATGACACT 1045  
 QY 341 AsnValPhePheLysGlnHisAspTyrIleIleGluPheAspTyrThrThrValLeu 360  
 DB 1046 AATGTGTCTTCAAGCAACATGATTAATCATCATTTGATGATTTTGTATTAATCTACTGCTTG 1105  
 QY 361 LysProSerLeuAsnPheThrAlaThrValLysValThrArgAlaAspGlyAsnGlnLeu 380  
 DB 1106 AAGCCATCTTCAACTTCACAGCCACTGTGATGATTAATCTGTCTGATGCAACCAACTG 1165  
 QY 381 ThrLeuGluGluArgArgAsnValIleIleThrValThrGlnArgAsnTyrThrGlu 400  
 DB 1166 ACTCTGAAGAAGAAAGAAATATATGATCATTAACGTGACACAGAGAACTATACGTAG 1225  
 QY 401 TyrTyrPheSerLysSerAsnSerGlyAsnGlnLysMetGluAlaValGlnLysIleAsnTyr 420  
 DB 1226 TACTGAGGCGGATCTAAACGTGAAATCAAGAAATGGAAGCTGTTCAAGAAATTAAT 1285  
 QY 421 ThrValProGlnSerGlyThrPheLysIleGluPheProIleLeuGluAspSerSerGlu 440  
 DB 1286 ACTGTCCCAAGATGAACTTTTAAAGTGAATTCCTCAATCTCTGAGGATTCAGAGTGA 1345  
 QY 441 LeuGlnLeuLysAlaTyrPheLeuGlySerLysSerSerMetAlaValHisSerLeuPhe 460  
 DB 1346 CTACAGTGAAGGCTTATTTCTTGTGATGATTAAGATGATGCAATGCAATCATCTGTTT 1405  
 QY 461 LysSerProSerLysThrTyrIleGlnLeuLysThrArgAspGluAsnIleLysValGly 480  
 DB 1406 AAGTCTCTATGATGAATATCAATCCAACTTAAACCAAGATGATGAAATATTAAGGTGGGA 1465

QY 481 SerProPheGluLeuValValSerGlyAsnLysArgLeuLeuGluLeuSerTyrMetVal 500  
 Db 1466 TGGCCTTTTGAAGTGGTGGTAGTGCAACAACGATTGAGAGATTGACTATATGTA 1525  
 QY 501 ValSerArgGlyGlnLeuValAlaValGlyGlnAsnSerThrMetPheSerLeuThr 520  
 Db 1526 GATATCCAGGGGACAGTGGTGGCTGTAGGAAAAACAATTCACAAATGTTCTCTTAAACA 1585  
 QY 521 ProGluAsnSerThrProLysValaCysValIleValTyrTyrIleGluAspArg 540  
 Db 1586 CCGAATAATTTCTGGAGCTCCAAAGCCTGTGATTTGTATTTATTTAAGATGATGGG 1645  
 QY 541 GlnIleIleSerAspValLeuLysIleProValGlnLeuValPheLysAsnLysIleLys 560  
 Db 1646 GAAATTATATAGTGAATGTTCTTAAAAAATTCCTGTACGCTTTTAAAAAATGAATAAAG 1705  
 QY 561 LeuTyrTrpSerLysValLysAlaGluProSerGluLysValSerLeuArgIleSerVal 580  
 Db 1706 CTATATTTGGAGTAAGTGAAGCTGAACCATCTGAGAAAAGTCTCTTGTGATCTCTGTG 1765  
 QY 581 ThrGlnProAspSerIleValGlyIleValAlaValAspLysSerValAsnLeuMetAsn 600  
 Db 1766 AACACGCTGACCTCCATAGTGGATTTGAGCTGTGACAAAAGTGTGAATCTGATGAAT 1825  
 QY 601 AlaSerAsnAspIleThrMetGluAsnValAlaIleGlnLeuGluLeuTyrAsnThrArg 620  
 Db 1826 GCCTCAATGATTTATTCATGAGAAAATGTGTCCAGACTTGGAACTTTATTAACACAGA 1885  
 QY 621 TyrTyrLeuGlyMetPheMetAsnSerPheAlaValPheGlnGluCysGlyLeuTrpVal 640  
 Db 1886 TATTTATTTGGCATTTTCATGAAATCTTTTGGACGTCTTCAGGAATGTGAGCTCGGGTA 1945  
 QY 641 LeuThrAspAlaAsnLeuThrLysAspTyrIleAspGlyValTyrAspAsnAlaGluTyr 660  
 Db 1946 TTGACAGATGCAAACTCAAGAAAGATTAATTTGATGATGATTAATACAAATCAGAAATAT 2005  
 QY 661 AlaGluArgPheMetGluGluAsnGluGlyIleValAspIleIleAspPheSerLeu 680  
 Db 2006 GGTGAGAGTTTATGAGAGAAAATGAGACACATTTGTGATATTCATCACTTTCTTTG 2065  
 QY 681 GlySerSerProHisValArgLysHisPheProGluTrpIleTrpLeuAspThrAsn 700  
 Db 2066 ATGTAGAGATCCACATGTCCGAAAGCATTTTCCAGACACTTGATTTGGCTAGACACCAAC 2125  
 QY 701 MetGlyTyrArgIleTyrGlnGlnPheGluValThrValProAspSerIleThrSerTrp 720  
 Db 2126 ATGGGTTTCAGAGATTTTACCAAGAAATTTGAAGTAACTGTACTGTATCTATCACTTCTGG 2185  
 QY 721 ValAlaThrGlyPheValIleSerGluAspLeuGlyLeuGlyLeuThrThrProVal 740  
 Db 2186 GTGGCTACGTGTTTGTATCTCTGAGGACCTGGGCTTGGACTTACCACTACCTCCAGTG 2245  
 QY 741 GlnLeuGlnAlaPheGlnProPhePheIlePheLeuAsnLeuProTyrSerValIleArg 760  
 Db 2246 GAGCTCCAGACCTTCCAACTATTTTCAATTTTGAATCTTCCACTCTGTATACGA 2305  
 QY 761 GlyGlnGluPheAlaLeuGluIleThrIlePheAsnTyrLeuLysAspAlaThrGluVal 780  
 Db 2306 GGTGAGAAATTTGCTTGGAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2365  
 QY 781 LysValIleIleGluLysSerAspLysPheAspIleLeuMetThrSerSerGluIleAsn 800  
 Db 2366 AAGGTATATATGAGAAAAGTGCACAAATTTGATTTCTATATGATCTTCAATATGAATTAAT 2425  
 QY 801 AlaThrGlyHisGlnGlnThrLeuLeuValProSerGluAspGlyAlaThrValLeuPhe 820  
 Db 2426 GCCACAGGCGCACAGACGACCTTGTGTCCAGTGAAGATGGGCGCACTGTCTTTT 2485  
 QY 821 ProIleArgProThrHisLeuGlyGluIleProIleThrValThrAlaLeuSerProThr 840  
 Db 2486 CCCATCAGGCGCAACATCTGGAGAAATTTCTATCAGTCAAGCTCTTTCACCCACT 2545

QY 841 AlaSerAspAlaValThrGlnMetIleLeuValLysAlaGluGlyIleGluLysSerTyr 860  
 Db 2546 CCTTCTGATGCTGTCAACCCACATGATTTTATGTAAGGCTGAAGGAATGAATAATCATAT 2605  
 QY 861 SerGlnSerIleLeuLeuAspLeuThrAspAsnArgLeuGlnSerThrLeuLysThrLeu 880  
 Db 2606 TCACATTCATCTTATTTAGACTTGACTGACAAATAGGCTACAGAGTACCTTAAAACTTTG 2665  
 QY 881 SerPheSerPheProProAsnThrValThrGlySerGluAspValGlnIleThrAlaIle 900  
 Db 2666 AGTTTCTATTTCTCTCAATACAGTCACTGGCAGTGAAGAAAGTTCAAGATCACTGCAAT 2725  
 QY 901 GlyAspValLeuGlyProSerIleAsnGlyLeuAlaSerLeuIleArgMetProTyrGly 920  
 Db 2726 GGAAGATGTTCTTGGTCTTCCATCAATGATGGCTTGAAGCTCATTTGATGGATGCTTATGGC 2785  
 QY 921 CysGlyGlnGlnAsnMetIleAsnPheAlaProAsnIleTyrIleLeuAspTyrLeuThr 940  
 Db 2786 TGTGTGAAACAGAACATGATTAATTTGCTCCAAATATTTTACATTTTGGATTATCTGACT 2845  
 QY 941 LysLysLysGlnLeuThrAspAsnLeuLysGluLysAlaLeuSerPheMetArgGlnGly 960  
 Db 2846 AAAAAGAAACAACTGACAGATTAATTTGAAGAAAAGCTTTTCAATTTATAGAGCAAGCT 2905  
 QY 961 TyrGlnArgGluLeuLeuTyrGlnArgGluAspGlySerPheSerAlaPheGlyAsnTyr 980  
 Db 2906 TACAGAGAGAACTTCTCTATACAGAGGAAATGGCTTTTCAAGTCTTTTGGGAATATAT 2965  
 QY 981 AspProSerGlySerThrTrpLeuSerAlaPheValLeuArgCysPheLeuGluAlaAsp 1000  
 Db 2966 GACCTTTCGGAGACACTGTTGTGTCAGCTTTTGTATTAAGATGTTTCTTGAAGCCGAT 3025  
 QY 1001 ProTyrIleAspIleAspGlnAsnValLeuHisArgThrTyrThrTrpLeuLysGlyHis 1020  
 Db 3026 CTTTACATAGATTTGATCAAGAAATGTGTTACACAGAAATACACTGTGCTTAAGAGCAT 3085  
 QY 1021 GlnLysSerAsnGlyGluPheTrpAspProGlyArgValIleHisSerGluLeuGlnGly 1040  
 Db 3086 CAGAAATCCAAACGGGATTTTGGATTCAGAGAAAGATTCATAGTACCTTCACAGCT 3145  
 QY 1041 GlyAsnLysSerProValThrLeuThrAlaTyrIleValThrSerLeuLeuGlyTyrArg 1060  
 Db 3146 GGCAATTAAGTCCAGTAACCTTAACCTATATTTGTAACCTTCTCTCGGGAATATAGA 3205  
 QY 1061 LysTyrGlnProAsnIleAspValGlnLysSerIleHisPheLeuGlnSerGluPheSer 1080  
 Db 3206 AAGTATCAGCCTTAACATTTGATGTGCAGAGCTTATCCATTTTGGAGCTGAATTCAGT 3265  
 QY 1081 ArgGlyIleSerAspAsnTyrThrLeuAlaLeuIleThrTyrAlaLeuSerSerValGly 1100  
 Db 3266 AGAGGAATTTACAGCAATTAATTAATCTTAACCTTATTAATTAATTAATTAATTAATTAAT 3325  
 QY 1101 SerProLysAlaLysGluAlaLeuAsnMetLeuThrTrpArgAlaGluGlnGluGly 1120  
 Db 3326 AGTCTTAAGGAGAAAGCTTTGAATATGCTGACTTGAAGACAGCAAGAAAGGTGGC 3385  
 QY 1121 MetGlnPheTrpValSerSerGluSerLysLeuSerAspSerTrpGlnProArgSerLeu 1140  
 Db 3386 ATGCAATTCGGGTGCTCATCAGAGTCCAAACTTTTCTGACTCTGTGGCAGCCAGCTCCCTG 3445  
 QY 1141 AspIleGluValAlaAlaTyrAlaLeuLeuSerHisPheLeuGlnPheGlnThrSerGlu 1160  
 Db 3446 GATATTAAGTTGAGCTTATGACCTGTCTTCAACTTCTTAACATTTTCAAGCTTCTGAG 3505  
 QY 1161 GlyIleProIleMetArgTrpLeuSerArgGlnArgAsnSerLeuGlyPheAlaSer 1180  
 Db 3506 GGAATCCCAATTAAGAGGTGGCTAAGCAGGCAAGAAATAGCTTGGGTGTTTTCATCT 3565  
 QY 1181 ThrGlnAspThrThrValAlaLeuLysAlaLeuSerGluPheAlaAlaLeuMetAsnThr 1200  
 Db 3566 ACTCAGGATACACGTGTGGCTTTTAAAGCTTGTCTGATTTTGAAGCCCTTATGAAATGCA 3625  
 QY 1201 GluArgThrAsnIleGlnValThrValThrGlyProSerSerProSerProLeuAlaVal 1220

```

Db      3626 GAAAGGACAAATATCCAGAGTACCGTGCAGGGGGCTGACCAAGCTCTTGCTGTG
Qy      1221 ValGlnProMetAlaValAsnIleSerAlaAsnGlyPheGlyPheAlaIleCysGluLeu
Db      3686 GTACAGCCACAGCGCATTAATATTCGCCAAATAGTGTTCATTTGATTCAGCTC
Qy      1241 AsnValValTyrAsnValIleSerGlySerSerArgArgArgSerIleGlnAsn
Db      3746 AATGTTGATATATGTGAAGCTTCCTGCTTCAAGAACAGAAATCTATCAAAAT
Qy      1261 GlnGluAlaPheAspLeuAspValAlaValIleGluLeuLeuAspAspLeuAsnHisVal
Db      3806 CAAAGAGCTTTGATTAGATGTTCGTAAAGAAATAAAGATGATCTCAATCATGTG
Qy      1281 AspLeuAsnValCysThrSerPheSerGlyProGlyArgSerGlyMetAlaLeuMetGlu
Db      3866 GATTGTAATGTGTGTCAAGCTTTCCGGCCCGGTAGAGTGGCCTTATATGGA
Qy      1301 ValAsnLeuLeuSerGlyPheMetValProSerGluAlaIleSerLeuSerGluThrVal
Db      3926 GTTAACCTATTAAAGTGTTCATGCTTATGCTTCAGAACAAATTTCTCTGAGCAGACAGTG
Qy      1321 LysIleValGluTyrAspHisGlyLysLeuAsnLeuTyrLeuAspSerValAsnGluThr
Db      3986 AAGAAAGGAAATGATCATGAAACTCAACCTCTATTAGATTCGTAAATGAAACC
Qy      1341 GlnPheCysValAsnIleProAlaValArgAsnIleValSerAsnThrGlnAspAla
Db      4046 CAGTTTGTGTAAATTCCTCGCTGTGAAACTTTAAAGTTTCAAAATCCCAAGATGCT
Qy      1361 SerValSerIleValAspTyrTyrGluProArgArgGlnAlaValArgSerTyrAsnSer
Db      4106 TCAGGTCCATAGTGATTAATAGAGCCAAAGAGACAGGGGTGAAGAAAGTTCAACTCT
Qy      1381 GluValIleLeuSerSerCysAspLeuCysSerAspValGlnGlyCysArgProCysGlu
Db      4166 GAACTGAAGCTGTCTCTCTGTGACCTTTGACGATGTCCAGGGCTGCGCTTGTGTAG
Qy      1401 AspGlyIleSerGlySerHisHisIleSerSerValIlePheIlePheCysPheIleLeu
Db      4226 GATGAGCTTCAAGCTCCCATCATCATCTTCAGTCATTATTTCTGTTCAGCTT
Qy      1421 LeuTyrPheMetGluLeuTyrLeu 1428
Db      4286 CTGTACTTATGGAACCTTGCGCTG 4309

RESULT 2
ID      AAL49815 standard; cDNA; 4335 BP.
XX      AAL49815;
XX      AC
XX      AAL49815;
XX      04-DEC-2002 (first entry)
XX      Human platelet alloantigen Gova coding sequence.
XX      Human: platelet alloantigen: Gova; Govb: single nucleotide polymorphism;
XX      SNr: biallelic; bleeding disorder; post-transfusion purpura;
XX      Km post-transfusion platelet refractoriness; haemostatic; vaccine;
XX      neonatal alloimmune thrombocytopenia; gene; ss.
XX      Homo sapiens.
XX      OS
XX      FH
XX      Key      Location/Qualifiers
XX      CDS      1..4335
XX      FT      /*tag= a
XX      FT      /product= "Gova"
XX      FT      /partial
XX      FT      /note= "no stop codon"
XX      FT      variation
XX      FT      replace(2108,C)
XX      FT      /*tag= b

```

```

PN      W0200270738-A2.
XX      12-SEP-2002.
XX      07-MAR-2002; 2002NO-CA00291.
XX      07-MAR-2001; 2001US-273941P.
XX      (SCHU//) SCHUH A.
XX      PI      Schuh A, Ouwehand W;
XX      WP1; 2002-713460/77.
XX      DR      P-PSDB; AAO19372.
XX      PT      New isolated oligonucleotide binding to a region of CD109 nucleic acid
XX      PT      having a single nucleotide polymorphism that distinguishes a Gova
XX      PT      and/or Govb allele, useful for treating blood disorders e.g. alloimmune
XX      PT      thrombocytopenia.
XX      PS      Claim 6; Page 23-29; 69pp; English.
XX      CC      The present invention relates to a sequence capable of binding
XX      CC      specifically to a CD109 nucleic acid which has a single nucleotide
XX      CC      polymorphism that distinguishes the Gova and Govb alleles. Detection of
XX      CC      the Gov genotype is useful for detecting whether the subject has or is at
XX      CC      risk of a blood disease, disorder or abnormal physical state, such as
XX      CC      bleeding, or increased risk of bleeding, due to alloimmune destruction of
XX      CC      blood platelets, e.g., post-transfusion purpura, post-transfusion
XX      CC      platelet refractoriness or neonatal alloimmune thrombocytopenia. The
XX      CC      nucleic acid and polypeptide are useful for Gov genotyping or phenotyping
XX      CC      individuals. The present sequence is the Gova coding sequence.
XX      SQ      Sequence 4335 BP; 1296 A; 837 C; 922 G; 1280 T; 0 other:

Alignment Scores:
Pred. No.:      0      Length:      4335
Score:          7326.50      Matches:      1427
Percent Similarity: 98.82%      Conservative: 1
Best Local Similarity: 98.75%      Mismatches: 0
Query Match:      99.71%      Indels:      17
DB:              24      Gaps:      1

US-10-020-095-4 (1-1428) x AAL49815 (1-4335)
Qy      1 MetGlnGlyProProLeuLeuThrAlaHisLeuLeuCysValCysThrAlaAlaLeu 20
Db      1 ATGCAAGGACCCACCGCTCTGACCGCGCCACCTCTCTGCGTGCACCGCGGCTG 60
Qy      21 AlaValAlaProGlyProArgPheLeuValThrAlaProGlyIleIleArgProGlyGly 40
Db      61 GCCGTGGCTCCGGGCGCTGGTTCTGTGTACAGCCCGAGGATCATCAGCGCGGAGGA 120
Qy      41 AsnValThrIleGlyValGluLeuLeuGluHisCysProSerGlnValThrValIleAla 60
Db      121 AATGTGCTATTGGGTGGAGCTTCGTGAACACTGCCCTTCACAGGTGACTGTGAAGCGG 180
Qy      61 GluLeuLeuYrThrIleSerAsnLeuThrValSerValLeuGluIleGluIleValPhe 80
Db      181 GACGTGTCAAGACAGATCAAACTCACTGTCTGTCTGTGAAGCAGAAAGATCTTT 240
Qy      81 GluIleGlySerPheIleYrThrLeuThrLeuProSerLeuProLeuAsnSerAlaAspGlu 100
Db      241 GAABAAGGCTCTTTAAAGACACTTACTCTTCATCACTACCTCTGAACAGTGCAATGAG 300
Qy      101 IleTyrGluLeuArgValThrGlyArgThrGlnAspGluIleLeuPheSerAsnSerThr 120
Db      301 ATTATAGAGCTACGTGAACCGGACCGATCCAGAGATGAGATTTATTTCTAATAGTACC 360
Qy      121 ArgLeuSerPheGluThrIleArgIleSerValPheIleGlnThrAspIleValIleTyr 140
Db      361 CGCTATCATTTGAGACCAAGAAATATCTGTCTTCAATTCACAGACAGAGGCTTATAC 420

```



OY	141	LYSP	ProLysGlnGluValLysPheArgLLeValThrLeuPheSerAspPheLysProCys	160
Db	421	AA	CCCAAGCAAGAGTAAGTTTCGCAATGTTACACTTCTCGAATTTAAAGCTTAC	480
OY	161	LYS	ThrSerLeuAsnLLeuLLeLysAspProLysSerAsnLeuLLeGlnLTrpLeu	180
Db	481	AAAA	CTCTTTAAACATTTCTCACTTAAGAGCCCAATTCAAATTTGATCCACAGGTG	540
OY	181	SerGln	InserAspLeuGlyValLLeSerLysThrPheGlnLeuSerSerHisProLle	200
Db	541	TCACA	CAAGAGATCTTGAGTCATTTCCAAACTTTTCAGCTATCTCCCATCCATA	600
OY	201	LeuGly	AspTrpSerLLeGlnValGlnValAsnAspGlnTrpTrpGlnSerPheGln	220
Db	601	CTTGAG	TGACTGCTCATTCACAGTTCAAGTAATGCGACACATATTCACATTCATTAG	660
OY	221	ValSerGln	TrpValLeuProLysPheGlnValThrLeuGlnThrProLeuTrpCysSer	240
Db	661	GTTTCAG	ATATGTTATACCAAAATTTGAAGACTTTGCAGACACCACTTATATTGTTCT	720
OY	241	MetAsn	SerLysHisLeuAsnGlyThrLLeThrAlaLysTrpThrTrpGlyLysProVal	260
Db	721	ATGAAT	CTTAAGCATTTTAATGGTACCATCAAGCGAAAGTATACATATGGGAAGCAGTG	780
OY	261	LysGly	AspValThrLeuThrPheLeuProLeuSerPheTrpGlyLysLysAsnLLe	280
Db	781	AAAGGAG	CGTAACCTTACATTTTACCTTATCTTTGGGGAAACAGAAATATTT	840
OY	281	ThrLys	ThrPheLysLLeAsnGlySerLLeAsnPheSerPheAsnAspGlnLMetCys	300
Db	841	ACAAAA	CAATTTAATTAATGGATCTGCAACTTCTCTTTAATGATGAAGAGATGAA	900
OY	301	AsnVal	MetAspSerSerAsnGlyLeuSerGlnTrpLeuAspLeuSerSerProGlyPro	320
Db	901	AATGTA	ATGGAATCTTCTCAATGGACTTTCTGATACCTGATCTATCTTCCCTGGACA	960
OY	321	ValGln	LLeuThrThrValThrGlnSerValThrGlyLLeSerArgAsnValSerThr	340
Db	961	GTAGA	AAATTTTAAACACAGTGCAGAAACAGTTACAGGATATTTCAAGAAATGAAGCACT	1020
OY	341	AsnVal	PhePheLysGlnHisAspTrpLLeLLeGlnPhePheAspTrpThrValLeu	360
Db	1021	AATGTT	CTCTTCACACATGATTAACATCTGATTTTGGATTATACACTACTGCTCTTG	1080
OY	361	LysPro	SerLeuAsnPheThrAlaThrValLysValThrArgAlaAspGlyAsnGlnLeu	380
Db	1081	AAGCAT	CTCTCAACTTCACAGCCACTGTGAAGATTAATCTCGTCTGATGGCACCACTG	1140
OY	381	ThrLeu	GlnGluArgAsnAsnValLLeThrValThrGlnArgAsnTrpThrGln	400
Db	1141	ACTCTT	GAAAGAAATTAATGTACTTAACAGTACACAGACAGAAACTATACTGAG	1200
OY	401	TyrTrp	SerGlySerAsnSerGlyAsnGlnLysMetGlnAlaValGlnLysLLeAsnTrp	420
Db	1201	TACTG	AGCGGATCTCAACAGTGAATACAGAAATGAGAAAGCTGTCCAGAAATTAATAT	1260
OY	421	ThrVal	ProGlnSerGlyThrPheLysLLeGlnPheProLleLeuGlnAspSerSerGln	440
Db	1261	ACTG	CCCCCAAGGGAACCTTTTAAGTTGAATTTCCCAATCTCGAGAGATTTCCAGTGG	1320
OY	441	LeuGln	LeuLysAlaTrpPheLeuGlySerLysSerSerMetAlaValHisSerLeuPhe	460
Db	1321	CTACAG	TGAAGGCTTATTTCTTGTTGTAATAAGTACATGGCAGTTCAATATGCTGTTT	1380
OY	461	LysSer	ProSerLysThrTrpLLeGlnLeuLysTrpTrpArgAspGluAsnLLeLysValGly	480
Db	1381	AAGT	TCTCTAGTAAACATACATCACTTAATAACAGAGTGAATAATTAAGGTGCGA	1440
OY	481	SerPro	PheGlnLeuValValSerGlyAsnLysArgLeuLysGlnLeuSerTrpMetVal	500
Db	1441	TGCG	CTTTTGAGTTGGTGGTTATGTGGCAACAAACGATTTGAAGAGATTTAAGCTATATGGTA	1500
OY	501	ValSer	ArgGlyGlnLeuValAlaValGlyLysGlnAsnSerTrpMetPheSerLeuThr	520

Db	1501	GTATCCAGGGACACTGGTGGCTGTAGGAAAACAAATTCACAACTGTTCTTTAA	156
QY	521	ProGluAenseTTrpThrProValaAcysValIleValaTyTrilIeGluAspAspGly	540
Db	1561	CCAGAAATTCCTTGACCTCCAAACCCCTGTGTAAATTCGTATTAATTCGAAGATGATGG	162
QY	541	GluIleIleSerAspValIleuLysIleProValGlnLeuValPheLysAnLysIleLys	560
Db	1621	GAAATTAATAGAGATGTTCTAAATAATTCCTGTTCAAGCTTGTTTTTAAATATAGATAAAG	168
QY	561	LeuTyTrpSerLysValIlysaIagIuProSerGluLysValSerLeuArgIleSerVal	580
Db	1661	CTAATTTGAGTAAGTAAGTAAGCTAAACCATCTGAGAAAGCTCTCTTGAAGCTCTGTG	174
QY	581	ThrgInProAspSerIleValGlyIleValAlaValAlaAspLysSerValAsnLeuMetAsn	600
Db	1741	ACACAGCCTGACTCCATAGTTCGAGTTGTAGCTGTGTGACAAAGTGTGAATCTGAAGAT	180
QY	601	AlaSerAsnAspIleThrMetGluAsnValAlaHisGluLeuGluLeuTyTrpAsnThrgIy	620
Db	1801	GCCCTTAATGATATTAACAATGGAAATATGGTCCATGAGTTGAACCTTATTAACACAGGA	186
QY	621	TyTrTyrlleuGlyMetPheMetAsnSerPheAlaValPheGlnGluCyseGlyLeuTPval	640
Db	1861	TATTATTTAGGATGATTCATGAAATCTTTTGACAGCTTTCAGAGAAATGTGACTCGGGTA	192
QY	641	LeuThrAspAlaAsnLeuThryLysAspTyTrIleAspGlyValTyTrAspAsnIagIuTy	660
Db	1921	TTGACAGATGCAAACCTCACAGAGATTAATTAATGATGGTGTTAAGACATGACAGAAAT	198
QY	661	AlaGluArgPheMetGluGluAsnGluGlyHisIleValAspIleHisAspPheSerLeu	680
Db	1981	GCTGAGAGCTTTATGAGAGAAATGAAGACATATGTGATATTCATGACTTTCCTTTG	204
QY	681	GlySerSerProHisValArgLysHisPheProGluThrTrpIleTrpLeuAspThrAsn	700
Db	2041	GGTACAGATCCACATGCTCGAAGACATTTTCCAGACACTGTGATTTGGCTAGACACCAAC	210
QY	701	MetGlyTyTrAspGlyIleTyTrGlnGluPheGluValIThrValProAspSerIleThrsertP	720
Db	2101	ATGGGTTACAGGATTTACACAGAAATTTGAAGTAAGTAAGTCTGATTCATACATTCCTGG	216
QY	721	ValAlaThrArgLysPheValIleSerGluAspLeuGlyIleuGlyLeuThrThrProVal	740
Db	2161	GTGGCTACTGGTTTTGTATCTCTAGGACCTGGGCTTGGACATAACAATCACTCCAGTG	222
QY	741	GluLeuGlnAlaPheGlnProPhePheIlePheLeuAsnLeuProTyrSerValIleArg	760
Db	2221	GAGCTCCAGGCTTCCAAACCAATTTTTCATTTTTCGAACTTCCCTACCTGTTATCAAG	228
QY	761	GlyGluGluPheAlaLeuGluIleThrIlePheAsnTyrlleuLysAspAlaThrGluVal	780
Db	2281	GGTGAAGAAATTTGCTTGGAAATACTATATTCATTAATTTTAAAGATCCCACTGAGGT	234
QY	781	LysValIleIleGluLysSerAspLysPheAspIleLeuMetThrserserGluIleAsn	800
Db	2341	AAAGTAATCATTTGAGAAAGTACAAATTTGATATCTTAATGACTTCAATATGAATAAAT	240
QY	801	AlaThrGlyHisGlnGlnThrLeuLeuValProSerGluAspGlyAlaThrValLeuPhe	820
Db	2401	GCCACAGGCCCACAGCAGACACCTCTGTGGTCCACAGTAGAGATGGGGCACTGTTCTTTT	246
QY	821	ProIleArgProThrHisIleuGlyGluIleProIleThrValThrAlaLeuSerProThr	840
Db	2461	CCCATCAGGCCAACACATCTGGGAAATAATCTTATCACAGTACACAGCTCTTTCACCCACT	252
QY	841	AlaSerAspAlaValThrgIuMetIleLeuValIlysaIagGlyTylIeGluLysSerTy	860
Db	2521	GCTTCTGATGCTGTCACCCAATGATTTTATGTAAGGCTGAAGGAATAGAAAAATCATAT	258
QY	861	SerGlnSerIleLeuLeuAspLeuThrAspAsnArgLeuGlnSerThrLeuTyThrLeu	880

D	2581	TCACATTCATCTTTTAACTTGACGTGACGACAAATAGGCTACAGAGTACCTGAAAACCTTTCG	2640
Q	881	SerPheSerPheProPheAntThrValThrGlySerGluArgValGlnIleThrAlaIle	900
D	2641	AGTTTCATCTTCTCTTAATACAGTGCACGTGAGGAAAGATTTCAGATCACTGCAAT	2700
Q	901	GlyAspValLeuGlyProSerIleAsnGlyLeuAlaSerLeuIleArgPheProTyrGly	920
D	2701	GGAATATCTTGTGCTCTTCATCAATAGGCTTACCTCATATGATGCGAATGCTTATGCG	2760
Q	921	CysGlyGlnGlnAsnMetIleAsnPheAlaProAsnIleTyrIleLeuAspTyrLeuThr	940
D	2761	TGTGTGAACAGAACAGATTAATTTTGTCCAAATATTTACATTTTGGATTTATCTGACT	2820
Q	941	LeuLeuValGlnLeuThrAspAsnLeuLeuGlyLeuValaLeuSerPheMetArgGlnGly	960
D	2821	AAAAAGAAACAACTGACAGATTAATTTGAAAGAAAAGCTCTTTCATTTATGAGCAAGT	2880
Q	961	TyrGlnArgGlnLeuLeuTyrGlnArgGlyLeuSerPheSerAlaPheIleValantYr	980
D	2881	TACACAGAGAACTTCTCTACAGAGGAGATGGCTCTTTCAGTCTTTTGGGAATTA	2940
Q	981	AspProSerGlySerThrTyrPheLeuSerAlaPheValLeuArgCysPheLeuGlnAlaAsp	1000
D	2941	GACCTTCTGGAGACATTTGCTGTGACGCTTTTGTTTTAAAGATTTTCTTGAACCCAT	3000
Q	1001	ProTyrIleAspIleAspGlnAsnValLeuAlaArgTyrTyrThrTyrPheLeuGlyAla	1020
D	3001	CTTACATATGATATGATGACGAATGTGTTTACAGAAACATCACTTGGCTTAAAGACAT	3060
Q	1021	GlnLeuSerAsnGlyGlnPheTyrPheProGlyValValIleHisSerGlyLeuGlnGly	1040
D	3061	CAGAAATCCAAACGTGATTTTGGGATCCAGAAAGATGATTCATGACGCTTCAAGT	3120
Q	1041	GlyAsnLeuSerProValThrLeuThrAlaTyrIleValThrSerLeuLeuGlyTyrArg	1060
D	3121	GCGATTAAGAGTCAGTAAACATTTACAGCTTATATGTAATCTCTCTGGGAATAGA	3180
Q	1061	LeuTyrGlnProAsnIleAspValGlnGlySerIleHisPheLeuGlnSerGlnPheSer	1080
D	3181	AAGATACAGGCTTAACATTTGATGCAAGAGTATCATTTTGGAGCTGAAATTCAGT	3240
Q	1081	ArgGlyIleSerAspAsnTyrThrLeuAlaLeuIleThrTyrAlaLeuSerValGly	1100
D	3241	AGAGGAATTTTCAGCAATTTACTCTTACGCCCTTAACTTAATGATTCATCACTGGGG	3300
Q	1101	SerProValAlaValGlnAlaLeuAsnMetLeuThrTyrArgAlaGlnGlnGlnGlyGly	1120
D	3301	AGTCTTAAGCGAAGAGGCTTGAATATGCTACTTGAAGGCAACAAAGAAAGTGGC	3360
Q	1121	MetGlnPheTyrValSerSerGlySerIleLeuSerAspSerTyrPheProArgSerLeu	1140
D	3361	ATGCAATTTCTGGGTGTCATCAGAGTCCAAACTTTCGACTCCCTGGAGCAAGCTCCCTG	3420
Q	1141	AspIleGlnValAlaAlaTyrAlaLeuLeuSerHisPheLeuGlnPheGlnThrSerGln	1160
D	3421	GATATTAAAGTTGACAGCTTATGACATGCTCTCACACTTCTTCAATTTTCAGACTTCTAG	3480
Q	1161	GlyIleLeuProLeuMetArgTyrPheLeuSerArgGlnArgAsnSerLeuGlyIlePheAlaSer	1180
D	3481	GGAATCCCAATTAAGAGGTGCTAAGCAGGCAAAAGAAATAGCTGGGTGGTTTTCATCT	3540
Q	1181	ThrGlnAspThrThrValAlaLeuValaLeuSerGlnPheAlaAlaLeuMetAsnThr	1200
D	3541	ACTCAGAGATACCACTGTGGCTTTAAAGGCTGTCTGATTTTGGACCCCTTAATGAATACA	3600
Q	1201	GlnArgThrAsnIleGlnValThrValThrGlyProSerSerProSerPro-----	1217
D	3601	GAAAGACAATAATCCAAAGTGAACGGTGAACGGGGCTACAGCTCACCAAGTCTGTAAAGTTT	3660
Q	1218	-----LeuAlaValValGlnPro	1223
D	3661	CTGATTGACACACAAACCGCTTATCTTTCAGACAGACAGAGAGCTGCTGTGTGATACGCCA	3720

Q	1224	MetAlaValAsnIleSerAlaAsnGlyPheGlyPheAlaIleCysGlnLeuAsnValVal	1243
D	3721	ATGGCAGTTAAATATTTCCGAATGTTTGGATTTGCTATTTTTCAGCTCAAGTGTGA	3780
Q	1244	TyrAsnValValaSerGlySerSerArgArgArgSerIleGlnAsnGlnAla	1263
D	3781	TATATGTGAAGCTTCTGGGCTTCTTGAAAGACCAAGATCTTATCCAAATCAAGAGCC	3840
Q	1264	PheAspLeuAspValAlaValGlyGlnAsnLysAspAspLeuAsnIleValAspLeuAsn	1283
D	3841	TTTGATTTGATGCTGCTGTAATAAAGATATGATCTCAATTCATGTGATTTGAT	3900
Q	1284	ValCysThrSerPheSerGlyProGlyArgSerGlyMetAlaLeuMetGlnValAsnLeu	1303
D	3901	GTTGTTCAAAGCTTTTGGGCGGGGTGAGAGTGGCATGCTCTTATGGAAGTTAACCTA	3960
Q	1304	LeuSerGlyPheMetValProSerGlnAlaIleSerLeuSerGluThrValValValVal	1323
D	3961	TTAAGTGTATATGATGCTTCAAGACCAATTTCTGACGCGAGACGTGAAGAAAGTG	4020
Q	1324	GluTyrAspHisGlyLeuLeuAsnLeuTyrLeuAspSerValaAsnGluThrGlnPheCys	1343
D	4021	GATATGATCAGGAAGAACTCACTTATTAAGTCTGTAAATGAAGCCAGTTTGT	4080
Q	1344	ValAsnIleProAlaValaArgAsnPheLysValSerAsnThrGlnAspAlaSerValSer	1363
D	4081	GTTAATATTCCTGCTGTGAGAACTTTAAATTTCAATACCAAGATGCTTCAGTCTCC	4140
Q	1364	IleValAspTyrTyrGlnProArgArgGlnAlaValaArgSerTyrAsnSerGlnValLys	1383
D	4141	ATAGTGATTTACTATGAGCAGGAGGAGGAGGCGGAGAAAGTTTCAACTGGAAGTAG	4200
Q	1384	LeuSerSerCysAspLeuLeuSerAspValGlnGlyCysArgProCysGlnAspGlyAla	1403
D	4201	CTGTCTCTGTGACTCTTGGAGTATGTCAGAGGCTCCCTCTTGTGAGGATGAGCT	4260
Q	1404	SerGlySerHisHisSerSerValIlePheIlePheCysPheLysLeuLeuTyrPhe	1423
D	4261	TCAGGCTCCCATCATCACTTCACTCAATTTTATTTTCTGTTCAAGCTTCTGTACTTT	4320
Q	1424	MetGluLeuThrLeu	1428
D	4321	ATGGAACCTTGGCTG	4335

RESULT 3

ABQ79964

ID ABQ79964 standard; cDNA; 4761 BP.

AC ABQ79964;

DT 23-DEC-2002 (first entry)

DE Human CD109 KI protein encoding cDNA.

KW CD109; alpha2 macroglobulin; cholesteryl ester; cerebroprotective; cardiant; immunosuppressive; haemostatic; anticoagulant; thrombolytic; human; cardiovascular; vasotropic; gene therapy; CD109 KI; gene; ss.

OS Homo sapiens.

PH Key

FT CDS

FT /tag= a

FT /product= "CD109 KI"

PN WO200270696-A2.

PD 12-SEP-2002.

PF 07-MAR-2002; 2002WO-CA00292.

XX 07-MAR-2001; 2001US-273814P.

PR

QY	1224	MechAlaValMetIleSerAlaAsnGlyPheGlyPheAlaIleCysGlnLeuAsnValVal	1243
Db	3721	ATGGCACTTAATATTTCCGCAAAATGGTTTGGATTTCCTATTTTGTCTGACGCTCAATGGTTGTA	3780
QY	1244	TyrAsnValIlysaIAsnGlySerSerSerArgArgArgSerIleGlnAsnGlnGluAla	1263
Db	3781	TATATATGAAAGGCTTCTGGGCTCTTCTTAAGAGACGAAGATCTATCCAAATATCAGAAAGCC	3840
QY	1264	PheAspLeuAspValAlaValIlyGlnAsnIlyAspAspLeuAsnIlyValAspLeuAsn	1283
Db	3841	TTTGATTATTAGATGTGTGCTGTAAAGAAATAAAGATCATCTCAATCATGTGGATTGGAAT	3900
QY	1284	ValCysThrSerPheSerGlyProGlyArgSerGlyMetAlaLeuMetGluValAsnLeu	1303
Db	3901	GTTGGTCAAGCTTTTCCGGCCCGGGTGAAGAGTGGCATGGCTCTTATGGAAAGTTAACCTTA	3960
QY	1304	LeuSerCylPheMetValProSerGlnAlaIleSerLeuSerGluThrValIlyIlyVal	1323
Db	3961	TTAAGTGGCTTTAAGTGCGCTTCTCAAGAGCAATTTCTCTGAGCAGACAGAGAAAGATG	4020
QY	1324	GluTyrAspHisGlyIlyIlyLeuAsnLeuTyrLeuAspSerValAsnGluThrGlnPheCys	1343
Db	4021	GAATATGATCATGGAAAACTCAACCTCTATTAGATTCTGTAAATGAACCAGTTTGT	4080
QY	1344	ValAsnIleProAlaValArgAsnPheIlyValSerAsnThrGlnAspAlaSerValSer	1363
Db	4081	GTTAATATTTCTGCTGTGAGAAACCTTTAAAGTTTCAAAATCCCAAGATGCTTCAAGTGTCC	4140
QY	1364	IleValAspTyrTyrGluProArgArgGlnAlaValArgSerTyrAsnSerGluValIlyLe	1383
Db	4141	ATATGTGATTACTATATGAGCCAGAGAGACAGCGGTGAGAAATTACCACTGTGAAGTGAAG	4200
QY	1384	LeuSerSerCysAspLeuCysSerAspValGlnGlyCysArgProCysGluAspGlyAla	1403
Db	4201	CTGTCCCTCTGTACCTTTTGCAAGATGTCCAGGCGCGCTCTTGTAGAGATGAGCT	4260
QY	1404	SerGlySerHisHisHisSerSerValIlePheIlePheCysPheIlyLeuLeuTyrPhe	1423
Db	4261	TCAGGCTCCCATCATCATCTCTTCAGTCATTTTATTTTCTGTTTCAAGCTTCTGTACTTT	4320
QY	1424	MetGluLeuThrLeu 1428	
Db	4321	ATGGAACTTGTGGCTG 4335	
RESULT 3			
ABQ79964	ID	ABQ79964 standard; cDNA; 4761 BP.	
XX	AC	ABQ79964;	
XX	DT	23-DEC-2002 (first entry)	
XX	DE	Human CD109 KI protein encoding cDNA.	
XX	KM	CD109; alpha2 macroglobulin; thioester; cerebroprotective; cardiant;	
XX	KW	immunosuppressive; haemostatic; anticoagulant; thrombolytic; human;	
XX	OS	cardiovascular; vasotropic; gene therapy; CD109 KI; gene; ss.	
XX	XX	Homo sapiens.	
XX	XX	Key	Location/Qualifiers
XX	XX	CDS	113..4450
XX	XX	FT	/*tag= a
XX	XX	FT	/product= "CD109 KI"
XX	XX	XX	W0200270696-A2.
XX	XX	XX	12-SEP-2002.
XX	XX	XX	07-MAR-2002; 2002W0-CA00292.
XX	XX	XX	07-MAR-2001; 2001US-273814P.
XX	XX	XX	PR

XX	(SCHU/) SCHUH A.
PA	(SOTH/) SOTHERLAND R D.
PA	Schuh A, Sutherland RD;
PI	WPl: 2002-713450/77.
DR	P-PStD; AB82165.
DR	New CD109 nucleic acids and polypeptides, useful in gene therapy,
XX	particularly for treating strokes, myocardial infarctions, thrombosis,
PT	thrombocytopenia, autoimmune diseases, or organ or bone marrow
PT	transplantation -
XX	
PS	Claim 1; Fig 1a; 156pp; English.
XX	
CC	The invention relates to isolated nucleic acid molecules encoding CD109
CC	polypeptides. These nucleic acid molecules include the human cDNA
CC	sequences comprising CD109 KI, CD109 KI-H7, CD109 K15 or their variants.
CC	CD109 is a novel member of the alpha2 macroglobulin (alpha2M)/C3, C4, C5
CC	family of cholesterol-containing proteins. The CD109 polypeptides can be
CC	expressed by standard recombinant methodology. The CD109 nucleic acid,
CC	CD109 polypeptide or its mimic, or the anti-CD109 antibody is useful
CC	for treating or detecting a disease or disorder, e.g. conditions
CC	associated with endothelial activation, platelet activation, activation
CC	of the coagulation or fibrinolytic systems, activation of T lymphocytes
CC	and of the complement system, quantitative or qualitative abnormalities
CC	of platelet function, increased or impaired platelet aggregation and
CC	activation, increased or impaired activation of the coagulation and/or
CC	fibrinolytic systems, or impaired or increased immune activation. These
CC	are also useful for treating cardiovascular disorders, stroke, myocardial
CC	infarction, thrombosis, embolism, peripheral vascular disease,
CC	thrombocytopenia, thrombocythemia, autoimmune diseases, organ
CC	transplantation, or bone marrow transplantation. The CD109 nucleic acids
CC	or their antisense nucleotide sequence are useful in gene therapy for
CC	treating these conditions. The present sequence represents the human
CC	CD109 KI cDNA sequence.
XX	
SQ	Sequence 4761 BP; 1420 A; 912 C; 1016 G; 1413 T; 0 other:
	Alignment Scores:
	Pred. No.: 0 Length: 4761
	Percent Similarity: 7326.50 Matches: 1427
	Best Local Similarity: 98.82% Conservative: 1
	Query Match: 99.71% Mismatches: 0
	Gaps: 1 Indels: 17
	Gaps: 24
US-10-020-095-4 (1-1428) x ABQ79964 (1-4761)	
OY	1 MetGlnGlyProProLeuLeuThrAlaAlaHisLeuLeuCyValCysThrAlaAlaLeu 20
Db	113 ATGAGGGGCCACCGCTCCGCAGCGCGGCCACTTCCTGTGGTGTGAACGCGCGCTG 177
OY	21 AlaValAlaProGlyProArgPheLeuValIThrAlaProGlyIleIleAryProGlyIy 40
Db	173 GCCTGTGCTCCCGGGCTCGCTTTCTGTGTGAACAGCCCAAGGATATATAGCCCGGAGA 233
OY	41 AsnValThrIleGlyValGluLeuLeuGluHisCysProSerGlnValIThrValIyAla 60
Db	233 AAATGTGACTATTGGGGGTGAGGCTTCGTGAAACCTGCCCTTCAAGTGACTGTAAGGCG 292
OY	61 GluLeuLeuLeuThrLaserAsnLeuThrValserValIleGluAlaGluGlyValPhe 80
Db	293 GAGCTGCTCAAGACAGCATNAACCTCACCTGCTCTGTCTGTGAAGAGGAAGAGTCTTT 352
OY	81 GluIySGlySerPheIyThrLeuIThrLeuProSerLeuProLeuAsnSerAlaSPGlu 100
Db	353 GAATAAGGCTCTTTTAAGACACTTACTCTTCATACACTACCTCTGAACAGTCAGTAGAG 412
OY	101 IleTyGluLeuValIThrGlyArGthrGlnASPGLuIleLeuPheSerAsnSerThr 120
Db	413 ATTATATAGCTACGTGAACCGGACCTAACCCAGATGAGATTATTTCTCTCATATAGTACC 472

QY	121	ArgLeuSerPheGluThrLysArgLLeSerValPheIleGlnThrAspLysAlaIleuYr	140
Db	473	CGCTTATCAATTGGAACCAAGAAATATCTGCTTCATTCACAAACAGACAGGCCCTTATAC	532
QY	141	LysProLysGlnGluValLysPheArgLLeValThrLeuPheSerAspPheLysProYr	160
Db	533	AAGCCAAACAGAAAGTGAAGTTGGCATTTGTACACTCTTCAGATTTTAAAGCTTAC	592
QY	161	LysThrSerLeuAsnIleLeuIleLysAspProLysSerAsnLeuIleGlnGlnTrpLeu	180
Db	593	AAAACCTCTTAAACATTCCTCATTAAGAGCCCAATCAAAATTTGATCCACAGGTGGT	652
QY	181	SerGlnGlnSerAspLeuGlyValIleSerLysThrPheGlnLeuSerSerHisProIle	200
Db	653	TCACAAACAAAGTGAAGCTTGGAGTCAATTTCCAAACCTTTTCAGCTATCTTCCATCCAAAT	712
QY	201	LeuGlyAspTrpSerIleGlnValGlnAlaAsnAspGlnThrTyThrGlnSerPheGln	220
Db	713	CTTGGTGACTGGTCTTATTCAGTTCAAGTGAAGTGCACAGACATATTAATCAATTTTCAG	772
QY	221	ValSerGlnTyThrValLeuProLysPheGlnValThrLeuGlnThrProLeuTyTrpCysSer	240
Db	773	GTTTCAGATATATGATTTACCAAAATTTGAAGTCACTTTGCAGACACCATTAATATGTTCT	832
QY	241	MetAsnSerLysHisLeuAsnGlyThrIleThrAlaLysTyThrTyThrGlyLysProVal	260
Db	833	ATGATATTTGAAGCATTTTAAATGATACCTACCGCAAAAGTATATACATATGCGAAACCACTG	892
QY	261	LysGlyAspValThrLeuThrPheLeuProLeuSerPheTrpGlyLysLysValAsnIle	280
Db	893	AAAGGAGACGTACCGCTTACATTTTACCTTTATCTTTGGGAAAGAGAAATATTT	952
QY	281	ThrLysThrPheLysIleAsnGlySerAlaAsnPheSerPheAsnAspGluLysMetCys	300
Db	953	ACAAATAATTTAAGATTAATATGATCTGCAAACTTCTTTTAAATGAATGAAGATGAATA	1011
QY	301	AsnValMetAspSerSerAsnGlyLeuSerGlnTyThrLeuAspLeuSerSerProGlyPro	320
Db	1013	AATGTAATGGAATTTCTTCAATGGACTTTCTGMAATACCTGGATCTATCTTCCCTGGACCA	1072
QY	321	ValGluIleLeuThrThrValIthrGluSerValThrGlyIleSerArgAsnValSerThr	340
Db	1073	GTAGAAATTTTAAACACAGTACAGAAACAGTTTCAAGGTATTTCAAGAAATGTAAGCACT	1133
QY	341	AsnValPhePheLysGlnHisAspTyrrIleIleGluPhePheAspTyThrThrValLeu	360
Db	1133	AATGGTCTTCAAGCAACATGATTAACATGATGATTTTGGATTAATATCATCTCTCTTG	1192
QY	361	LysProSerLeuAsnPheThrAlaThrValLysValThrArgAlaAspGlyAsnGlnLeu	380
Db	1193	AAGCATCTCTCAACTTCAACGCCACTGTGAAGGTAATCTCGTGTGATGGCAACCACTG	1252
QY	381	ThrLeuGlnGluArgArgAsnAsnValIleThrValIthrGlnArgAsnTyThrGlu	400
Db	1253	ACTCTTGAAGAAAGAAATTAATGTAAGTCAATTAACAGTACACAGAGAAACTATCTGAG	1312
QY	401	TyrTrpSerGlySerAsnSerGlyAsnGlnTyPheMetGluAlaValGlnHisIleAsnTy	420
Db	1313	TACTGAGACGGATCTCAACAGTGAATAGAAATAGAAAGAGCGTTTCAGAAATAATTAATAT	1372
QY	421	ThrValProGlnSerGlyThrPheLysIleGluPheProIleLeuGluAspSerSerGlu	440
Db	1373	ACTGTCCCCCAAAGTGAACCTTTTAAGTTGAATTTCCAAATCTTGAGAGATTTCCAGTGA	1433
QY	441	LeuGlnLeuLysAlaTyrrPheLeuGlySerLysSerSerMetAlaValHisSerLeuPhe	460
Db	1433	CTACAGTTGAAGGCTATTTCTTGGTGTGTAAGTAACATGACATGGCAGTTCAATGTCTGTTT	1492
QY	461	LysSerProSerLysThrTyrrIleGlnLeuLysThrArgAspGluAsnIleLysValGly	480
Db	1493	AAGTCTCTTAATGAACATATCATCCAAATTAATAACAAAGATGTAAATATTTAAAGGTGGGA	1552

```
Qy 481 SerProhegIuLeuValIaSerGIyASnIySgIuLeuIySgIuLeuSerTyMetVal 500
Db 1553 TCGCCTTTTGAAGTGGTGTAGTGGCAACAACGATTGAAGAGTTAAGCTATATGATA 1612
Qy 501 ValSerArgGIyGInLeuValIaValGIyIySgInaSerThrMetPheSerLeuThr 520
Db 1613 GTATCCAGGGAGCAGTGGTGGCTGTAGAAAACAATAATCAACAAGTCTCTTTAAACA 1672
Qy 521 ProGUaNSerPTThrProLySAlaCySValIleValTyTYrIleGUaNSpGIy 540
Db 1673 CCAAGAAATTTTGGACCTCCAAAAGCCGTGTAATGTGTATATTAAGAAAGATGGG 1732
Qy 541 GluIleIleSerAspValIleuLySleProValGInLeuValPheLySAsnLyIleLyS 560
Db 1733 GAAATTATTAAGTATGTTCTTAAATAATCTCTGTACGTTGTTTAAATAATTAAGATAAG 1792
Qy 561 LeuTYrTPSerLyValIyAlaGluProSerGIyLySValSerLeuArgIleSerVal 580
Db 1793 CTATATTTGAGTAAAGTAAAGCTGAACCATCTGAGAAAGTCTCTTAAGATCTGTG 1852
Qy 581 ThrGInProAspSerIleValGIyIleValIaValAspLySAspValAsnLeuMetAsn 600
Db 1853 ACAACGCTGACTCATAGTGGATGTGGATTTAGCTTTGACAAAAGTGTCAATCTGATGAT 1912
Qy 601 AlaSerAsnAspIleThrMetGluAsnValaHisGInLeuGluLeuTyAsnThrGIy 620
Db 1913 GCGCTTATATGATATTAACAATGGAATAATGTGTCATGAGTTGGAACCTTATACACAGA 1972
Qy 621 TyTYrIleuGIyMetPheMetAsnSerPheAlaValPheGInGluCySgIyLeuTYrVal 640
Db 1973 TATTATTTAGGACATGTCATGAATTCCTTTGACGCTTTGACGAAATGGAATCTGGGTA 2032
Qy 641 LeuThrAspAlaAsnLeuThrLySAspTYrIleAspGIyValTyAspAsnAlaGIyTy 660
Db 2033 TTGACAGATGCAAACTTCACAAAGSATTATATGTAGTGTATATGACAAATGCGAATAT 2092
Qy 661 AlaGluArgPheMetGluGluAsnGluGIyHisIleValAspIleHisAspPheSerLeu 680
Db 2093 GCTGAGGTTTATGAGGAAATGAAGGACATATTTGATATTAATTAATCACTTTTCTTTG 2152
Qy 681 GlySerSerProHisValArgLySHisPheProGluThrTYrIleTPLeuAspThrAsn 700
Db 2153 GGTAGCAGTCCACATGTCGAAAGCATTTTCCAGAGACTTGGAATTTGGCTAGACACCAAC 2212
Qy 701 MetGIyTYrArgIleTYrGInGluPheGluValThrValProAspSerIleThrSerTy 720
Db 2213 ATGGGATTACAGGATTACCAAGAAATTTGAAGTAATGTAACCTGATCTATATCACTTTGG 2272
Qy 721 ValAlaThrGIyPheValIleSerGIuAspLeuGIyLeuGIyLeuThrThrProVal 740
Db 2273 GTGGCTACTGGTTTGTGATCTCTGAGAGCCTGGGCTTGGACTTAACAACACTCACTG 2332
Qy 741 GluLeuGInAlaPheGInProPhePheIlePheLeuAsnLeuProTYrSerValIleArg 760
Db 2333 GAGCTCCCAAGCCTTCCAAACCATTTTTCATTTTGTGAATCTCCCTACTCTGTATGAGA 2392
Qy 761 GIyGIuGIuPheAlaLeuGIuIleThrIlePheAsnTYrLeuLySAspAlaThrGIuVal 780
Db 2393 GGTAAAGAAATTTGCTTGGAAATACATATTAATTAATTAATTAAGAGCCACGAGGTT 2452
Qy 781 LySValIleIleGIuLySAspLySAspPheAspIleLeuMetThrSerSerGIuIleAsn 800
Db 2453 AAGGTATTCATTGAGAAAGTGAACAATTTGATATTTCAATGACTTCAAAAGAAATTAAT 2512
Qy 801 AlaThrGIyHisGInGInThrLeuLeuValProSerGIuAspGIyAlaThrValLeuPhe 820
Db 2513 GCCACAGGCCACACGACACCTTCTGTGTTCCAGTGGAGATGGGCAACTGTTCTTTT 2572
Qy 821 ProIleArgProThrHisLeuGIyGluIleProIleThrValIleThrAlaLeuSerProThr 840
Db 2573 CCCATCCAGGCCAACATCTGGAGAAATTTCTTATCAAGTCAACAGCTTTTCACTCACT 2632
Qy 841 AlaSerAspAlaValThrGInMetIleuValLySAlaGIyIleGIuLySAspTy 860
Db 2633 GCTTCTGATGCTGCTCAACAGATGATTTAGTAAAGCTGCAAGAAATAGAAATCATAT 2692
Qy 861 SerGInSerIleLeuLeuAspLeuThrAspAsnArgLeuGInSerThrLeuIySThrLeu 880
Db 2693 TCAATATCATCTTATTAAGCTTGAAGTACAAATAGCTTACAGATACCTGAAAACCTTTG 2752
Qy 881 SerPheSerPheProAsnThrValThrGIySerGIuArgValGInIleThrAlaIle 900
Db 2753 AGTTTCTCATTTCTCTTAATACATGATCTGGCAGTGAAGAAAGTCAAGATCACTGCAAT 2812
Qy 901 GIyAspValLeuGIyProSerIleAsnGIyLeuAlaSerLeuIleArgMetProTYrGIy 920
Db 2813 GGAATGTTCTTGTGCTTCCATCAATGAGCTTGAAGCTTCATGATTTGGAGCTTTATGGC 2872
Qy 921 CySgIyGluGInaMetIleAsnPheAlaProAsnIleTYrIleLeuAspTYrLeuThr 940
Db 2873 TGTGTGAACAGAAACATGATTAATTTTGCTCCAAATTTTAATTTGATTTATCTGACT 2932
Qy 941 LySLeLySgInLeuThrAspAsnLeuLySgIuLyAlaLeuSerPheMetArgInGIy 960
Db 2933 AAAAAGAAACACTGACAGATTAATTTGAAGAAAAGCTCTTCATTTATGAGCAAGT 2992
Qy 961 TyGIuArgGIuLeuLeuTYrGInArgGIuAspGIySerPheSerAlaPheGIyAsnTYr 980
Db 2993 TACCAGAGAAACTTCTATCAGAGGGAAGATGGCTTTTCAGTCTTTTGGAAATTAAT 3052
Qy 981 AspProSerGIySerThrTPLeuSerAlaPheValLeuArgYsPheLeuGIuAlaAsp 1000
Db 3053 GACCTTCTGGAGCACTGTGTGTGACGCTTTGTTTAAGATGTTCTCTTGAAGCGGAT 3112
Qy 1001 ProTYrIleAspIleAspGInaValLeuHisArgTYrTYrThrTPLeuLySgIyHis 1020
Db 3113 CTTTACATGATATTTGATGATGATGTGTTTACAGAAACATACCTTGCTTAAGGACAT 3172
Qy 1021 GInLySAsnGInGluPheTPAspProGIyArgValIleHisSerGIuLeuGInGIy 1040
Db 3173 CAGAAATCCAAACGCTGATTTGGGATCCAGAGAGTGAATCATAGTGAAGCTTCAAGGT 3232
Qy 1041 GIyAsnLySAspProValThrLeuThrAlaTYrIleValHisSerLeuLeuGIyTYrArg 1060
Db 3233 GGCATTAATAAGTCCAGTAACCTTACAGCTTAATTTGTAATTTCTCTCCGGGATATAGA 3292
Qy 1061 LyS TYrGInProAsnIleAspValGInGluSerIleHisPheLeuGInSerGIuPheSer 1080
Db 3293 AAGTATACGCTTAACATGATGTGCAAGATCATCAATTTTGTGAAGTGTGAATCAAT 3352
Qy 1081 ArgGIyIleSerAspAsnTYrThrLeuAlaLeuIleThrTYrAlaLeuSerSerValGIy 1100
Db 3353 AGAGSAAATTCAGACAAATTAATTAAGCTTACAGCTTATTAATTAATTAATTAATTAAT 3412
Qy 1101 SerProLySAlaLySgIuAlaLeuAsnMetLeuThrTPArgAlaGIuGInGInGIy 1120
Db 3413 AGTCTTAAGCGAAGAAAGCTTTGAATATGTCTACTTGAGACAGCAAGAAAGGAGG 3472
Qy 1121 MetGInPheTYrValSerSerGIuSerLySLeuSerAspSerTPGInProArgSerLeu 1140
Db 3473 ATGCAATTCGGGTGTCATAGAGTCCAAACTTTTGAATCTTCCGCGACCAACCTCCCTG 3532
Qy 1141 AspIleGIuValAlaAlaTYrAlaLeuLeuSerHisPheLeuGInPheGInThrSerGIu 1160
Db 3533 GATATTGAATGTCAGGCTTACAGCTGCTCACACTTTCAATTTCAATTTCAAGCTTCTGAG 3592
Qy 1161 GIyIleProIleMetArgTPLeuSerArgGInArgAsnSerLeuGIyGIyPheAlaSer 1180
Db 3593 GGAATCCCAATTTAGGTGTGCTTAACAGGCAAGAAATAGCTTGGGTGTGTTTGAATCT 3652
Qy 1181 ThrGInAspThrThrValAlaLeuLySAlaLeuSerGIuPheAlaAlaLeuMetAsnThr 1200
Db 3653 ACTCAGGATCCACTGTGCTTTAAAGCTCTGTCTGAATTTGCAAGCCCTTAATGAATACA 3712
Qy 1201 GluArgThrAsnIleGInValThrValThrGIyProSerSerProSerPro----- 1217
```





Db 593 AAAACCTCTTAAACATTCCTAATGAAGACCCCAATCAATTGATCCACAGCTGGTTG 652  
Qy 181 SerGlnGlnSerAspLeuGlyValIleSerIleThrPheGlnLeuSerSerHisProIle 200  
Db 653 TCACAAACAAAGTGTCTTGAGTCAATTCCTCAAACTTTTCAGCTATCTTCCCATCCATA 712  
Qy 201 LeuGlyAspTyrSerIleGlnValGlnValAsnAspGlnThrTyrTyrGlnSerPheGln 220  
Db 713 CTGGTGCCTGCTCATTCAGATTCAGATGAAATGACGACGACATATTCATTCATTTCCAG 772  
Qy 221 ValSerGlyTyrValLeuProIlePheGlnValIleThrLeuGlnThrProLeuTyrCysSer 240  
Db 773 GTTTCAGAAATGATGATTAACCAAAATTTGAAAGTGAATGACCAACCAATATATTTGTTCT 832  
Qy 241 MetAsnSerIleSerHisLeuAsnGlyThrIleThrAlaIleThrTyrTyrGlyLeuProVal 260  
Db 833 ATGAAATTCATAGCATTTTAATGTAACATTCAGGCAAAATGATACATATGGAACCCAGTG 892  
Qy 261 LysGlyAspValIleThrLeuThrPheLeuProLeuSerPheTyrGlyValValAsnIle 280  
Db 893 AAAGAGACGTAACGCTTACATTTTACCTTATCTTTGGGAAAGAAATAATATT 952  
Qy 281 ThrIleThrPheIleValIleAsnGlySerAlaAsnPheSerPheAsnAspGlyGlnMetLys 300  
Db 953 ACAAACCAATTAAGATTAATGATCTGCAAACTTCTTTTAATGATGAAGAGATGAAA 1012  
Qy 301 AsnValMetAspSerSerAsnGlyLeuSerGlyTyrTyrLeuAspLeuSerSerProGlyPro 320  
Db 1013 AATTTAATGATTTCTTCAATGACCTTCTGAAATCTGATCTGATCTTCCCCGAGAACCA 1072  
Qy 321 ValGlyIleLeuThrThrValIleThrGlySerValIleThrGlyIleSerArgAsnValSerThr 340  
Db 1073 GTACAAATTTTAACCAAGTACAGAAATCAGTTACAGATTTTCAAGAAATGTAAGACCT 1132  
Qy 341 AsnValIlePheIleValIleAsnGlySerAlaAsnPheSerPheAsnAspGlyGlnMetLys 360  
Db 1133 AATGTGTCTTCAAGCAACATGATTAATCATGATTTTGAATTTAATCACTACGTCCTG 1192  
Qy 361 LysProSerLeuAsnPheThrAlaThrValIleValIleThrArgAlaAspGlyAsnGlnLeu 380  
Db 1193 AAGCATCTCTCAACTTACAGCCACTGTGAAGGTAACTCGTGCAGAGGCAACCAACTG 1252  
Qy 381 ThrLeuGlnGlyAsnArgAsnAsnValIleThrValIleThrGlnArgAsnTyrThrGln 400  
Db 1253 ACTTTGAAGAAAGAAATTAATGTAAGTCAATACAGTACACAGAGAACTTAATACGAG 1312  
Qy 401 TyrTrpSerGlySerAsnSerGlyAsnGlnIleMetGlnValIleGlnIleValIleAsnTyr 420  
Db 1313 TACTGAGCCGATCTTAACAGTGAATCAAGAAATGGAAGCTGTTCAGAAATTAATATAT 1372  
Qy 421 ThrValIleProGlnSerGlyThrPheIleIleGlnPheProIleLeuGlnAspSerSerGln 440  
Db 1373 ACTGCCCCCAAGGGAACCTTTTAAGTGAATTTCCAAATCCGAGAGATTTCCAGTAG 1432  
Qy 441 LeuGlnLeuIleValAlaTyrPheLeuGlySerIleSerSerMetAlaValIleSerIlePhe 460  
Db 1433 CTACAGTTGAAGGCTAATTTCTTGTAAGTGAAGTCAAGGATTCATAGCTGTTT 1492  
Qy 461 LysSerProSerIleThrTyrIleGlnLeuIleThrArgAspGlnAsnIleValIleValIle 480  
Db 1493 AAGTCTCTTACTAAGACATACATCACTAAACCAAGAGATGAAGAAATTAATGAAGTGGCA 1552  
Qy 481 SerProPheGlnLeuValIleSerGlyAsnIleValIleValIleValIleValIleVal 500  
Db 1553 TCGCTTTTGAAGTGTGTTAGTGAACAAACCAATGAAGAGTTAAGCTAATATGTA 1612  
Qy 501 ValSerArgGlyGlnLeuValAlaValIleValIleValIleValIleValIleValIleVal 520  
Db 1613 GTATCCAGGAGACAGTGTGCTGCTGAGAAACAAATTCACAAATTTCTCTTAAACA 1672  
Qy 521 ProGlnAsnSerTyrProIleProIleValIleValIleValIleValIleValIleValIle 540  
Db 1673 CCAAAATATCTTGAGACTCAAAAGCTGTAAATGTGATTAATGAAGATGATGGG 1732

Qy 541 GlnIleIleSerAspValIleLeuValIleProValIleLeuValIlePheValIleValIleVal 560  
Db 1733 GAATATTAAGTATGTTCTTAATAAATTTCCGTTCACTGCTTTTAAATAAATGAATGAAG 1792  
Qy 561 LeuTyrTrpSerIleValIleValIleValIleValIleValIleValIleValIleValIle 580  
Db 1793 CTATATTTGAGTAAAGTGAAGGCTGAACCATCTGAGAAATGCTCTCTTACGATCTGCTG 1852  
Qy 581 ThrGlnProAspSerIleValIleValIleValIleValIleValIleValIleValIleVal 600  
Db 1853 ACACAGCTTACCTCAGTATGAGATTTGATGCTGTGACAAAGTGAATGAAATCTGAATAT 1912  
Qy 601 AlaSerAsnAspIleThrMetGlnAsnValIleValIleValIleValIleValIleValIle 620  
Db 1913 GCCTCTATATATTTCAATGAAATGCTGCTCATGATGATGGAATTTTAACACAGCA 1972  
Qy 621 TyrTyrLeuGlyMetPheMetAsnSerPheAlaValIlePheGlnGlyCysGlyLeuTyrVal 640  
Db 1973 TATTAATTTAGCAGTTCATGAATTTCTTTGACAGTCTTTTCAAGAAATGTCAGCTCTGGTA 2032  
Qy 641 LeuThrAspAlaAsnLeuThrIleValIleValIleValIleValIleValIleValIleVal 660  
Db 2033 TTGACAGATCAAACTCAGAGAGATTAATGATGCTTTATGACAAATGACAAATAT 2092  
Qy 661 AlaGlyArgPheMetGlnGlnGlnGlyIleIleValIleValIleValIleValIleValIle 680  
Db 2093 GCTGAGAGCTTTTGAAGAAATGAAGACATATTTGATTAATTCATGACTTTTCTTTG 2152  
Qy 681 GlySerSerProHisValIleArgIleValIlePheProGlnThrTyrIleThrLeuAspThrAsn 700  
Db 2153 GGTAGCAGTCCACATGCTCCAAAGCAATTTTCCAGACCTTGATTTGGCTAGACACCAAC 2212  
Qy 701 MetGlyTyrArgIleTyrGlnGlnIlePheGlnValIleValIleValIleValIleValIle 720  
Db 2213 ATGGGTACAGGATTTACAAAGATTTGAAGTACGTAATCTTATCATCCTCTTGG 2272  
Qy 721 ValAlaThrGlyPheValIleSerGlnAspLeuGlyLeuGlyLeuThrThrProValIle 740  
Db 2273 GTGGCTACTGTTTGTGATCTCTGAGACCTGGCTTGTGACTTAACATCACTCCAGTG 2332  
Qy 741 GlnLeuGlnAlaPheGlnProPhePheIlePheLeuAsnLeuProTyrSerValIleArg 760  
Db 2333 GAGCTCCAACTTCCAAACCATTTTCAATTTTGAATTTTCCCTCACTCTGTTATAGA 2392  
Qy 761 GlyGlnGlnPheAlaLeuGlnIleThrIlePheAsnTyrLeuIleValIleValIleValIle 780  
Db 2393 GGTGAAGAAATTTCTTTGAAATTAATCAATTAATTTGAAGATGACACTGAGGT 2452  
Qy 781 IleValIleIleGlnIleSerAspIlePheAspIleLeuMetThrSerSerGlnIleAsn 800  
Db 2453 AAGGTATCATTTGAGAAATGACCAATTTGATTTCTAATGACTTCAATGAATGAATTAAT 2512  
Qy 801 AlaThrGlyHisGlnGlnThrLeuLeuValIleProSerGlnAspGlyAlaThrValIlePhe 820  
Db 2513 GCCACAGGCCACAGCAGACCTTCTGCTTCCAGTGAAGATGGGCAACTGTTCTTTT 2572  
Qy 821 ProIleArgProThrHisLeuGlyGlnIleProIleThrValIleThrAlaLeuSerProThr 840  
Db 2573 CCATCAAGGCAACATCTGGAGAAATTCATATCACTCACTCACTCTTCACTCCACT 2632  
Qy 841 AlaSerAspAlaValIleThrGlnMetIleLeuValIleValIleValIleValIleValIle 860  
Db 2633 GCTTCTATCTGCTGACCCAGATGATTTTGTAAAGCTTAAGAAATGAATTAATCATAT 2692  
Qy 861 SerGlnSerIleLeuLeuAspLeuThrAspAsnArgIleGlnSerThrLeuIleValIle 880  
Db 2693 TCACAAATCAATTTAATGACTGACCAATAGGCTTACAGATCACTGAAACCTTGG 2752  
Qy 881 SerPheSerPheProProAsnThrValIleThrGlySerGlnArgValIleGlnIleThrAlaIle 900  
Db 2753 AGTTTCATATTTCTCTTAATACAGTACGAGTGAAGGAGTTGAGATCACTGCAATT 2812

QY 901 GlyAspValLeuGlyProSerIleAsnGlyLeuAlaSerLeuIleArgMetProTyrGly 920  
 Db 2813 GGAGAGTTCCTGGTCCCTTCATCATAGCTTCAATGCTGATGCGATGCTTATGCG 2872  
 QY 921 CysGlyGluGlnAsnMetIleAsnPheAlaProAsnIleTyrIleuAspTyrLeuThr 940  
 Db 2873 TGTGTGAAACAACAATGATTAATTTTGCTCAAAATATTACATTTTGATATATCTGACT 2932  
 QY 941 LysIleValGlnLeuThrAspAsnLeuLysGluValAlaLeuSerPheMetArgGlnGly 960  
 Db 2933 AAAAAGAAACAACAGTGCATGATATATTGAAGAAAAAGCTCTTTCATTATGAGGCAAGCT 2992  
 QY 961 TyrGlnAsnGluLeuLeuTyrGlnIleAspGlyLysPheSerAlaPheGlyAsnTyr 980  
 Db 2993 TACCGAGAGAACTTCTCTATACAGAGGAGATGCTCTTTCAGTCTTTTGGAATATAT 3052  
 QY 981 AspProSerGlySerThrTyrPheLeuSerAlaPheValLeuArgCysPheLeuGluAlaAsp 1000  
 Db 3053 GACCCCTTCGGAGAGCACTGGTGTGTCAGCTTTGTTTAAGATGTTTCTTGAAGCCGAT 3112  
 QY 1001 ProTyrIleAspIleAspGlnAsnValLeuHisArgThrTyrThrTyrLeuLysGlyHis 1020  
 Db 3113 CCTTCATATGATATATGATCATGATATGTTTACACAGAACATACACTTGGCTTAAAGACAT 3172  
 QY 1021 GlnLysSerAsnGlyValPheThrAspProGlyArgValIleHisSerGluLeuGlnGly 1040  
 Db 3173 CAGAAATCCAAACGGTGAATTTTGGGATCCAGAGAGATGATTCATATGATGACTTCAAGCT 3232  
 QY 1041 GlyAsnLysSerProValThrLeuThrAlaTyrIleValThrSerLeuLeuGlyTyrArg 1060  
 Db 3233 GGCATATAAAGTCCAGATACACTTACAGCTTATATGTAATCTCTCTCGGATATATA 3292  
 QY 1061 LysTyrGlnProAsnIleAspValGlnGlnSerIleHisPheLeuGlnSerGluPheSer 1080  
 Db 3293 AAGTATACCTTACCTTACCTTATATGTCAGAGCTCTATCATTTTGGAGCTGATTCAGT 3352  
 QY 1081 ArgGlyIleSerAspAsnTyrThrLeuAlaLeuIleThrTyrAlaLeuSerValGly 1100  
 Db 3353 AGAGGAATTTTCAGACATTAATATCTAGCCCTTATACCTTATGATTCATCATGAGTGGG 3412  
 QY 1101 SerProLysValLysGluAlaLeuAsnMetLeuThrTyrArgAlaGluGlnGluGly 1120  
 Db 3413 ACTCCTAAAGCCAAAGAGAGCTTGAATATGCTGACCTTGAGAGCAGAACAGAGCTGCG 3472  
 QY 1121 MetGlnPheTyrValSerSerGlyLysSerLysSerAspSerThrGlnProArgSerLeu 1140  
 Db 3473 ATGCAATTCCTGGGTGTCATCAAGTCCAAACTTTCGACTCCCTGGAGCAGCGCTCCCTG 3532  
 QY 1141 AspIleGluValAlaAlaTyrAlaLeuLeuSerHisPheLeuGlnPheGlnThrSerGlu 1160  
 Db 3533 GATATTTGAAGTTGACAGCTTATGACACTGCTCTCACACTTCTTCAATTTTCAGACTTCTGAG 3592  
 QY 1161 GlyIleProIleMetCysArgTyrPheLysSerArgGlnIleAsnSerLeuGlyGlyPheAlaSer 1180  
 Db 3593 GGAATCCCAATATATGAGTGGCTTAAGCAGGCAAAATATAGCTTGGGTGTTTGCATCT 3652  
 QY 1181 ThrGlnAspThrThrValAlaLeuLysValAlaLeuSerGlnPheAlaAlaLeuMetAsnThr 1200  
 Db 3653 ACTCGAGATACACAGTGGCTTAAAGGCTCTGTCTGATTTTTCAGAGCTTAAATGATAC 3712  
 QY 1201 GluArgThrAsnIleGlnValThrValThrGlyProSerSerProSerPro----- 1217  
 Db 3713 GAAAGGACAAATATCCAGATGACCGTGAAGGCGCTTACCTCAAGTCTCTGTAAGTTT 3772  
 QY 1218 -----LeuAlaValAlaGlnPro 1223  
 Db 3773 CTGATTTGACACACAAACGCTTACTCTTCAAGACAGACAGACTTGTGTACAGCCA 3832  
 QY 1224 MetAlaValAsnIleSerAlaAsnGlyPheGlyPheAlaIleCysGlnLeuAsnValVal 1243  
 Db 3833 ATGGCAGTTAATATATTCGCAATATGTTTGGATTTTGTGATTTGCTCATATGTTTGA 3892  
 QY 1244 TyrAsnValLysAlaSerGlySerSerArgArgArgSerIleGlnAsnGlnGluAla 1263

Db 3893 TATATGTAAGCTTCTGGGCTTCTTGTAGAACACCAAGATCTATCCAAATCAAGAGCC 3952  
 QY 1264 PheAspLeuAspValAlaValLysGlnLysAsnAspAspLeuAsnHisValAlaLeuAsn 1283  
 Db 3953 TTTGATTTAGATGTTGCTGTAAAGAAATTAAGATGATCTCAATCAATGATGATTTGAAT 4012  
 QY 1284 ValCysThrSerPheSerGlyProGlyArgSerGlyMetAlaLeuMetGluValAsnLeu 1303  
 Db 4013 GTGTGTAACAGCTTTTCGGGCCGGGTGAGATGSCATGCGCTTATGGAATTAACCTTA 4072  
 QY 1304 LeuSerGlyPheMetValProSerGluAlaIleSerLeuSerGluThrValLysIleVal 1323  
 Db 4073 TTAAAGTGGCTTATATGCTCTTCAAGACCAATTTCTGAGCGGACAGTGAAGAAAGTG 4132  
 QY 1324 GluTyrAspHisGlyLysLeuAsnLeuTyrLeuAspSerValAsnGluThrGlnPheCys 1343  
 Db 4133 GAATATGATCATGAGAAACCTCAACCTCTATTTAGATCTGTAAATGAACCCAGTTTGT 4192  
 QY 1344 ValAsnIleProAlaValArgAsnPheLysValSerAsnThrGlnAspAlaSerValSer 1363  
 Db 4193 GTTATATTTCTGCTGTGAGAACTTTAAAGTTTCAATACCCAGATCTTCAATGATTC 4252  
 QY 1364 IleValAspTyrTyrGluProArgArgGlnAlaValArgSerTyrAsnSerGluValLys 1383  
 Db 4253 ATATGGAATTACTATGAGCCAAAGAGACAGGCGGTGAAGATTAACAATCTGAAGTGAAG 4312  
 QY 1384 LeuSerSerCysAspLeuCysSerAspValGlnGlyCysArgProCysGluAspGlyVal 1403  
 Db 4313 CTGTCTCTCTGACCTTGTGCGATGATGTCCAGGCGCTCGCTCTTGTGAGGATGAGCT 4372  
 QY 1404 SerGlySerHisHisSerSerValIlePheIlePheCysPheLysLeuLeuTyrPhe 1423  
 Db 4373 TCAAGCTCCATCATCATCTTTCAGTCAATTTTATTTTCTGTTTCAAGCTTCTGATCTT 4432  
 QY 1424 MetGluLeuTyrPhe 1428  
 Db 4433 ATGGAATTTGGCTG 4447  
 Db 4433 ATGGAATTTGGCTG 4447  
 RESULT 5  
 ID ABQ79966  
 ABQ79966 etandard; cDNA; 5895 BP.  
 XX  
 AC ABQ79966;  
 XX  
 DT 23-DEC-2002 (first entry)  
 XX  
 DE Human CD109 K1-H7 protein encoding cDNA.  
 XX  
 KW CD109; alpha2 macroglobulin; cholesteryl; cerebroprotective; cardiac;  
 KW immunosuppressive; haemostatic; anticoagulant; thrombolytic; human;  
 KW cardiovascular; vasotropic; gene therapy; CD109 K1-H7; gene; ss.  
 OS Homo sapiens.  
 XX  
 FH Key location/Qualifiers  
 FT CDS 113..4450  
 FT /tag= a  
 FT /product= "CD109 K1-H7"  
 PN  
 PD WO200270696-A2.  
 XX  
 XX 12-SEP-2002.  
 XX  
 PF 07-MAR-2002; 2002WO-CA00292.  
 XX  
 PR 07-MAR-2001; 2001US-273814P.  
 XX  
 PA (SCHU/) SCHU A.  
 PA (SUTH/) SUTHERLAND R D.  
 XX  
 PI Schuh A, Sutherland RD;  
 XX

DR WPI: 2002-713450/77.  
DR P-PSDB; ABB82167.  
XX New CD109 nucleic acids and polypeptides, useful in gene therapy,  
PT particularly for treating strokes, myocardial infarctions, thrombosis,  
PT thrombocytopenia, autoimmune diseases, or organ or bone marrow  
PT transplantation -  
XX  
PS Claim 1; Fig 2a; 156pp; English.  
XX  
CC The invention relates to isolated nucleic acid molecules encoding CD109  
CC polypeptides. These nucleic acid molecules include the human cDNA  
CC sequences comprising CD109 K1, CD109 K1-H7, CD109 K15 or their variants.  
CC CD109 is a novel member of the alpha2 macroglobulin (alpha2M)/C3, C4, C5  
CC family of thioester-containing proteins. The CD109 polypeptides can be  
CC expressed by standard recombinant methodology. The CD109 nucleic acid  
CC CD109 polypeptide or its mimetic, or the anti-CD109 antibody is useful  
CC for treating or detecting a disease or disorder, e.g. conditions  
CC associated with endothelial activation, platelet activation, activation  
CC of the coagulation or fibrinolytic systems, activation of T lymphocytes  
CC and of the complement system, quantitative or qualitative abnormalities  
CC of platelet function, increased or impaired platelet aggregation and/or  
CC activation, increased or impaired activation of the coagulation and/or  
CC fibrinolytic systems, or impaired or increased immune activation. These  
CC are also useful for treating cardiovascular disorders, stroke, myocardial  
CC infarction, thrombosis, embolism, peripheral vascular disease,  
CC thrombocytopenia, thrombocythemia, autoimmune diseases, organ  
CC transplantation, or bone marrow transplantation. The CD109 nucleic acids  
CC or their antisense nucleotide sequence are useful in gene therapy for  
CC treating these conditions. The present sequence represents the human  
CC CD109 K1-H7 cDNA sequence.  
XX  
SQ Sequence 5895 BP; 1709 A; 1173 C; 1241 G; 1772 T; 0 other;

Alignment Scores:  
Pred. No.: 0 Length: 5895  
Score: 7326.50 Matches: 1427  
Percent Similarity: 98.82% Conservative: 1  
Best Local Similarity: 98.75% Mismatches: 0  
Query Match: 99.71% Indels: 17  
DB: 24 Gaps: 1

US-10-020-095-4 (1-1428) x ABO79966 (1-5895)

QY 1 MetGlnGlyProProLeuLeuThrAlaAlaHisLeuLeuCyValCysThrAlaLeu 20  
DB 113 ATGCAAGGCCCAACGCTCTTAACGCGCCCACTCTCTGCGTGAACGCGCGCTG 172  
QY 21 AlaValAlaProGlyProArgPheLeuValThrAlaProGlyIleIleArgProGlyGly 40  
DB 173 GCCCGGCTCCCGGCGCTCGGTTCTGGTGAACGCCCAAGGATCATAGGCCGAGGA 232  
QY 41 AsnValThrIleGlyValGlnLeuLeuGlnHisCysProSerGlnValThrValIleAla 60  
DB 233 AATGACTATTTGGGGTGGAGCTTCTCGAACAATGCCCTTCACAGGTACGTGACGCG 292  
QY 61 GlnLeuLeuLeuThrAlaSerAsnLeuThrValSerValLeuGlnAlaGlnGlyValPhe 80  
DB 293 GAGCGCTCAAGACACGATCAAACTCACTGCTCTGCTCGAAGCAGAAAGAGTCTTT 352  
QY 81 GlnIleGlySerPheLeuThrLeuThrLeuProSerLeuProLeuAsnSerAlaAspGlu 100  
DB 353 GAAAGAGGCTCTTTTAACACACTTACTTTCATCATCACTCTGGAACAGTGCAGATGAG 412  
QY 101 IleTyrGlnLeuValThrGlyArgThrGlnAspGlnIleLeuPheSerAsnSerThr 120  
DB 413 ATTATGAGTACGCTAAGCGAAGTACCCAGAGTGAATTTATTTCTTAATAGTACC 472  
QY 121 ArgLeuSerPheGlnThrIleArgIleSerValPheIleGlnThrAspValAlaLeuTyr 140  
DB 473 CGCTTATCATTTGAGACCAAGAAATATCTGCTTCATTCAAAAGAACGAGCCCTTATAC 532  
QY 141 LysProGlyGlnGlnValIleValPheArgIleValThrLeuPheSerAspPheLeuProTyr 160

DB 533 AAGCCAAAGCAAGAGGAAATTGCGATTGTACACTCTTCAGATTTTAAGCTTAC 592  
QY 161 LysThrSerLeuAsnIleLeuIleValAspProLysSerAsnLeuIleGlnGlnTyrLeu 180  
DB 593 AAAACCTCTTAAACATTCATTTAAGAACCCCAATCAAAATTTGATTCACAGCTGGTG 652  
QY 181 SerGlnIleSerAspLeuGlyValIleSerLysThrPheGlnLeuSerSerHisProIle 200  
DB 653 TCACAAACAAAGTATCTTGGAGTCATTTCCAAAACCTTTACGATATCTTCCATCCATA 712  
QY 201 LeuGlyAspTyrPheIleGlnValGlnValAsnAspGlnThrTyrTyrGlnSerPheGln 220  
DB 713 CTGGTACCTGCTCTATTCAGTTCAAGTCAAGACCAAGACATATTCATCATCTTTCAG 772  
QY 221 ValSerGlnTyrValLeuProLysPheGlnValThrLeuGlnThrProLeuTyrCysSer 240  
DB 773 GTTTCAGAAATATGATTAACCAAAATTTGAAGTGCTTGGACAGACCATATATGTTCT 832  
QY 241 MetAsnSerLysHisLeuAsnGlyThrIleThrAlaLysTyrThrTyrGlyLysProVal 260  
DB 833 ATGAATTTCAACATTTTAATGATGACATCAGCCAAAGTATCATATGGGAAGCCAGTG 892  
QY 261 LysGlyAspValThrLeuThrPheLeuProLeuSerPheTyrGlyLysLysAsnIle 280  
DB 893 AAGAGACGTTAACGCTTACATTTTACCTTTATCTTTGGGAAAGAAAGAAATATTT 952  
QY 281 ThrLysThrPheLysIleAsnGlySerAlaAsnPheSerPheAsnAspGlnGlnMetLys 300  
DB 953 ACAAACCAATTTAAGATTAATGATGATCGCAAACTCTCTTTAATGATGAAGATATAA 1012  
QY 301 AsnValMetAspSerSerAsnGlyLysSerGlnTyrLeuAspLeuSerSerProGlyPro 320  
DB 1013 AATGTATGATTTCTTCAATATGACCTTTCGAATACCTGGATCATCTTCCCGGACCA 1072  
QY 321 ValGlnIleLeuThrThrValThrGlnSerValThrGlyIleSerArgAsnValSerThr 340  
DB 1073 GTAGAAATTTTAAACACAGTGAACAGATCATGTTACAGTATTTCAAAATATGTAAGACT 1132  
QY 341 AsnValPhePheLysGlnHisAspTyrIleIleGlnPhePheAspTyrThrThrValLeu 360  
DB 1133 AATGTGTTCTTCAAGCAACATGATTCATCATGATTTTGTATTAATCACTAGCTCTTG 1192  
QY 361 LysProSerLeuAsnPheThrAlaThrValIleValThrArgAlaAspGlyAsnGlnLeu 380  
DB 1193 AAGCATCTCTCAACTTCACAGCCACCTGTAAGGTAACTGTGTATGGCAACCACTG 1252  
QY 381 ThrLeuGlnGlnArgArgAsnAsnValIleThrValThrGlnArgAsnTyrThrGln 400  
DB 1253 ACTCTGAAGAAAGAAATTAATGTAGTCATTAACATGACACAGAGAAACTATACGAG 1312  
QY 401 TyrTyrPheSerGlySerAsnSerGlyAsnGlnLysMetClnAlaValGlnLysIleAsnTyr 420  
DB 1313 TACTGAGCGGATCTTAACATGAGAAATCAAGAAATGGAACCTGTCGAAAAATTAATTA 1372  
QY 421 ThrValProGlnSerGlyThrPheLysIleGlnPheProIleLeuGlnAspSerSerGln 440  
DB 1373 ACTGTCCCCCAAGTGAACCTTTAAGATTTGAATTCCTCGAGAGATTCACAGTACG 1432  
QY 441 LeuGlnLeuLysAlaTyrPheLeuGlySerLysSerSerMetAlaValHisSerLeuPhe 460  
DB 1433 CTACAGTTGAAGGCTATTTCTTGTGTAATAAGTATGAGCATGGAGTTACATGCTGTTT 1492  
QY 461 LysSerProSerLysTyrTyrIleGlnLeuLeuThrArgAspGlnAsnIleValGly 480  
DB 1493 AAGTCTCTAGTATGACATCATCTCACTAAAAACAAGATGAATAATATAAAGTGGGA 1552  
QY 481 SerProPheGlnLeuValIleSerGlyAsnLysArgLeuLysGlnLeuSerTyrMetVal 500  
DB 1553 TCGCTTTTGAAGTTGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 1612  
QY 501 ValSerArgGlnLeuValAlaValGlyLysGlnAsnSerThrMetPheSerLeuThr 520

Db 1613 GATTCACGGGACAGTTGGCTGTAGAGAAAACAAATGTTCTCTTAACA 1672  
Qy 521 ProGluAsnSerTyrThrProLysAlaCysValIleValTyrTyrIleGluAspArgIy 540  
Db 1673 CCAGAAATATCTTGGACTCCAAAGCCTGTGTATTTGTATATTAATGAAGTATGGG 1732  
Qy 541 GluIleIleSerAspValLeuLysIleProValGluLeuValPheLysAsnLysIleLys 560  
Db 1733 GAATATATAGTATGATCTTCAAAAATCTCTGTTCACTGTTTAAAAATAGATAAAG 1792  
Qy 561 LeuTyrTyrSerLysValIleLysAlaGluProSerGluLysValSerLeuArgIleSerVal 580  
Db 1793 CTATATGGAGTAAAGTAAAGCTGAACATCTGAAGAAAGTCTCTTGAAGATCTGTGG 1852  
Qy 581 ThrGluProAspSerIleValGlyIleValAlaValAspLysSerValAsnLeuMetAsn 600  
Db 1853 AACAGCCTGACTCCCATAGTTGGATTTGAGCTGTGGACAAAAGTGTGATCTGATGAAT 1912  
Qy 601 AlaSerAsnAspIleThrMetGluAsnValIleHisGluLeuGluLeuTyrAsnThrGly 620  
Db 1913 GCCTCAATGATATTCATTAAGAAATGTGTCTCATGAGTGGAACTTTATACACAGGA 1972  
Qy 621 TyrTyrLeuGlyMetPheMetAsnSerPheAlaValPheGluGluCysGlyLeuTyrVal 640  
Db 1973 TATATTTAGGCAATGTTCAATGAAATCTTGTGAGTCTTTCAGAAATGTGACTGTGGTA 2032  
Qy 641 LeuThrAspAlaAsnLeuThrLysAspTyrIleAspGlyValTyrAspAsnAlaGluTyr 660  
Db 2033 TTGACAGATGCAAACCTTCAGAAAGATTTATATTGATGATGATTTATACAAATGCAAAAT 2092  
Qy 661 AlaGluArgPheMetGluGluAsnGluGlyHisIleValAspIleHisAspPheSerLeu 680  
Db 2093 GCTGAGAGGTTATAGAGAAATAGAAAGACATATTGTAGATATTCAATCATCTTCTTTGG 2152  
Qy 681 GlySerSerProHisValArgLysHisPheProGluThrTyrIleTyrLeuAspThrAsn 700  
Db 2153 GGATAGAGTCCACATGTCCGAAAGCATTTTCCAGACACTGTGATTTGGCTAGACACAAAC 2212  
Qy 701 MetGlyTyrArgIleTyrGluGluPheGluValIleThrValProAspSerIleThrSerTyr 720  
Db 2213 ATGGGTTACAGATTTTACCAAGAAATTTGAAGTAACTGTACCTGATTTCTATCACTTCTTGG 2272  
Qy 721 ValAlaThrGlyPheValIleSerGluAspLeuGlyLeuGlyLeuThrThrThrProVal 740  
Db 2273 GGGGACTACGTTTGTCTGATCTCTGAGACCTGGGCTTGGACTAACCACTCACTCCAGAG 2332  
Qy 741 GluLeuGluAlaPheGluProPhePheIlePheLeuAsnLeuProTyrSerValIleArg 760  
Db 2333 GAGCTCCAAAGCCTTCCAAACATTTTTCATTTTTCATCTTCCCTACTCTGTATACAGA 2392  
Qy 761 GlyGluGluPheAlaLeuGluIleThrIlePheAsnTyrLeuLysAspAlaThrGluVal 780  
Db 2393 GGGTAAAGAAATTTGCTTGGAAATTAATCAATTAATTTTGAAGAAATCCACTGAGGTT 2452  
Qy 781 LysValIleIleGluLysSerAspLysPheAspIleLeuMetThrSerSerGluIleAsn 800  
Db 2453 AAGGTATCATTTGAGAAAGTAGCAAAATTTGATATTCTATATGACTTCAATGAATGAATAT 2512  
Qy 801 AlaThrGlyHisGluGluIleThrLeuLeuValProSerGluAspGlyAlaThrValLeuPhe 820  
Db 2513 GCCACAGGCGCACAGACAGACCTTCTGTGTCCAGAGTAGGAGCACTGTCTCTTTT 2572  
Qy 821 ProIleArgProThrHisIleuGlyGluIleProIleThrValIleuSerProThr 840  
Db 2573 CCATATAGGCACACATCTGGAGAAATTCATATACAGTCAAGCTCTTTCACCCACT 2632  
Qy 841 AlaSerAspAlaValThrGluMetIleLeuValIleGluGlyIleGluLysSerTyr 860  
Db 2633 GCTTCGATGCTGTCAACCAAGATTTTATAGTAAAGCTGAAGGAATATAAAATATATAT 2692  
Qy 861 SerGlnSerIleLeuLeuAspLeuThrAspAsnArgLeuGlnSerThrLeuLysThrLeu 880  
Db 2693 TCACATCATCTTATATAGATTGACTGACAAATAGGCTACAGAGTACCTCGAAAACTTGG 2752

Qy 881 SerPheSerPheProAsnThrValThrGlySerGluArgValGluIleThrAlaIle 900  
Db 2753 AGTTTCTATTTCCCTCAATACAGTGAAGTGAAGGATTCAGATCATCTCAATT 2812  
Qy 901 GlyAspValLeuGlyProSerIleAsnGlyLeuAlaSerLeuIleArgMetProTyrGly 920  
Db 2813 GGAGATGTTCTTGTGCTCTTCCATCAATGAGCTTACCTCATGTGATTCGAGTGCCTATGGC 2872  
Qy 921 CysGlyGluGluAsnMetIleAsnPheAlaProAsnIleTyrIleLeuAspTyrLeuThr 940  
Db 2873 TGTGTGAACAGAAATATATTAATTTTCTCCAAATATTTTTCATTTTGGATTTATCTGACT 2932  
Qy 941 LysLysLysGluLeuThrAspAsnLeuLysGluLysAlaLeuSerPheMetArgGlnGly 960  
Db 2933 AAAAGAAACAACATGACAGATTAATTTGAAGAAAAAGCTCTTTCATTTATAGGCAAGT 2992  
Qy 961 TyrGluArgGluLeuLeuTyrGluArgGluAspGlySerPheSerAlaPheGlyAsnTyr 980  
Db 2993 TACAGAGAGAACTTCTCATACAGAGGAGATGGCTTTTCAGTGTCTTTGGAAATTAAT 3052  
Qy 981 AspProSerGlySerThrTyrTyrLeuSerAlaPheValLeuArgCysPheLeuGluAlaAsp 1000  
Db 3053 GACCTTCTGGAGACATTTGT 3112  
Qy 1001 ProTyrIleAspIleAspGluAsnValIleuHisArgThrTyrThrTyrTyrLeuLysGlyHis 1020  
Db 3113 CCTTACATAGATTTGATCAGAAATGTTTTACAGAAACATACATCTTGGCTTTAAAGACAT 3172  
Qy 1021 GlnLysSerAsnGlyGluPheTyrAspProGlyArgValIleHisSerGluLeuGlnGly 1040  
Db 3173 CAGAAATCCAAAGGATTTTGGATTTTGGATTTTGGATTTTGGATTTTGGATTTTGGATTTTGG 3232  
Qy 1041 GlyAsnLysSerProValIleThrLeuThrAlaTyrIleValIleThrSerLeuLeuGlyTyrArg 1060  
Db 3233 GGCATTAAGTCCAGTAACTTACACCTTATATTTGAACTTCTCTCTGGGATATGGA 3292  
Qy 1061 LysTyrGluProAsnIleAspValGluGluSerIleHisPheLeuGluSerGluPheSer 1080  
Db 3293 AAGTATTCAGCTTACATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3352  
Qy 1081 ArgGlyIleSerAspAsnTyrThrLeuAlaLeuIleThrTyrAlaLeuSerSerValGly 1100  
Db 3353 AAGGAATTTTCAAGAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3412  
Qy 1101 SerProLysAlaLysGluAlaLeuAsnMetLeuThrTyrArgAlaGluGlnGluGlyGly 1120  
Db 3413 AGTCTTAAAGCAGAGGAAGCTTTGAAATGTGCACTTGGAGAGCAAGAACAAAGGTGGC 3472  
Qy 1121 MetGlnPheTyrValSerSerGluSerLysLeuSerAspSerTyrGlnProArgSerLeu 1140  
Db 3473 ATGCATTTCTGGGTCTATCAGAGTCCAAACTTCTGTGCTCTGTGACCAACGCTCCCTG 3532  
Qy 1141 AspIleGluValAlaAlaTyrAlaLeuLeuSerHisPheLeuGluPheGluIleThrSerGlu 1160  
Db 3533 GATATTTGAAGTTGACGCTTAAGCACTGCTTCAACATCTTACATTTCAAGATCTTGAAG 3592  
Qy 1161 GlyIleProIleMetArgTyrTyrLeuSerArgGlnArgAsnSerLeuGlyPheAlaSer 1180  
Db 3593 GGAATCCCAATTAATGAGGTGTGTAAGCAGCAAGAAATTAACCTGGGCTGTGTGTGTGTGTGT 3652  
Qy 1181 ThrGluAspThrThrValAlaLeuLysValAlaLeuSerGluPheAlaAlaLeuMetAsnThr 1200  
Db 3653 ACTCAGATACCACTGTGTGTAAAGCTGTGTGAATTTGACGCCCTTAATGAATAC 3712  
Qy 1201 GluArgThrAsnIleGluValIleThrValThrGlyProSerSerProSerPro----- 1217  
Db 3713 GAAAGAGCAATATTCAGATGACGTGACGGGCTTACCAAGTCTGTAAAGTTT 3772  
Qy 1218 -----LeuAlaValIleGlnPro 1223  
Db 3773 CTGATTTGACACACAAACCGTTACTCTTCAAGACAGACAGAGCTTGTGTGTGTGTGTGTGTGT 3832

QY 1224 MetAlaValAsnIleSerAlaAsnGlyPheAlaIleCysGlnLeuAsnVal 1243  
Db 3833 ATGGCAGTTAATATTTCCGCAAAATGGTTTGGATTGCTATTGTCAGCTCAATGTTGTA 3892  
QY 1244 TyrAsnValLysAlaSerGlySerSerArgArgArgAspSerIleGlnAsnGlnAla 1253  
Db 3893 TATAATGGAAGGCTTCTGGGCTTCTTGAAGAACCAAGATCTATCCAAATCAAGAGCC 3952  
QY 1264 PheAspLeuAspValAlaValLysGlnAsnLysAspAspLeuAsnHsValAspLeuAsn 1283  
Db 3953 TTTGATTAGATGTTGCTGTAAGAAATAAAGATGATCTCAATCAGTGGATTGAT 4012  
QY 1284 ValCysThrSerPheSerGlyProGlyArgSerGlyMetAlaLeuMetGlnValAsnLeu 1303  
Db 4013 GTGTGTCACAGCTTTCCGGGCCCGGGTGTAGAGTGGCATGCTCTTATGGAAGTTAACCTA 4072  
QY 1304 LeuSerGlyPheMetValProSerGlnAlaIleSerLeuSerGlyThrValLysVal 1323  
Db 4073 TTAAGTGGCTTTATGATGCTTTCAGAACCAATTTCTTGAGCGAGACGTGAAGAAAGTG 4132  
QY 1324 GlnTyrAspHsIleGlyLysLeuAsnLeuTyrLeuAspSerValAsnGlnThrGlnPheCys 1343  
Db 4133 GAATATGATCATGGAACCACTCACTTATTAATTTGATGTAATGAACCACTTTTGT 4192  
QY 1344 ValAsnIleProAlaValArgAsnPheLysValSerAsnThrGlnAspAlaSerValSer 1363  
Db 4193 GTTAATATTTCCCTGCTGTGAGAACTTTAAAGTTTCAATATCCCAAGATGCTTCAGTGTCC 4252  
QY 1364 IleValAspIlyTyrGlnProArgArgGlnAlaValAspSerTyrAsnSerGlnValLys 1383  
Db 4253 ATAGTGATTAATCTAGACCAAGACAGACGCGGTGAAGAACTTCAACTGGAAGTGAAG 4312  
QY 1384 LeuSerSerCysAspLeuCysSerAspValGlnGlyCysArgProCysGlnAspGlyAla 1403  
Db 4313 CTGTCCTCTCTGACCTTGCAGTATGTCAGGGCTCCCTCTTGAGAGATGAGACT 4372  
QY 1404 SerGlySerHsIleHsIleSerSerValIlePheIlePheCysPheLysLeuLeuTyrPhe 1423  
Db 4373 TCAGGCTCCCATCATCACTTTCAGTCAATTTTATTTCTGTTCAAGCTTCTGACTTT 4432  
QY 1424 MetGlnLeuTyrPhe 1428  
Db 4433 ATGGAACCTTGGCTG 4447

RESULT 6  
AAD49435  
ID AAD49435 standard; DNA; 5883 BP.  
XX  
AC AAD49435;  
XX  
DT 24-MAR-2003 (first entry)  
XX  
DE Human r150 DNA #1.  
XX  
XX Human, GPI-anchored TGF-beta1 binding protein; r150 protein; cancer;  
KW glycosylphosphatidyl inositol; transforming growth factor-beta1;  
KW therapy; gene; de.  
XX  
OS Homo sapiens.  
XX  
XX Key Location/Qualifiers  
FH 113..4450  
CDS /\*tag=a  
FT /product= "Human r150 protein #2"  
FT /transl\_except= (pos:2219..2221, aa:Xaa)  
FT /note= "Xaa corresponds to Ser, Tyr"  
XX  
XX WO200285942-A2.  
XX  
XX 31-OCT-2002.  
XX  
XX 24-APR-2002; 2002WO-CA00560.  
XX

PR 24-APR-2001; 2001US-285713P.  
PR 14-FEB-2002; 2002US-356163P.  
XX  
PA (UYMC-) UNIV MCGILL.  
XX  
PI Philip A, Tam B;  
XX  
DR WPI; 2003-093100/08.  
XX P-PSDB; AAE32013.  
XX  
PT Novel transforming growth factor (TGF)-beta 1 binding reagent which  
PT comprises r150 protein which acts as accessory receptor of TGF-beta,  
PT useful for negatively modulating TGF-beta activity, and thus for  
PT treating cancer -  
XX  
PS Claim 7; Page 100-103; 127pp; English.  
XX  
CC The invention relates to novel transforming growth factor (TGF)-beta1  
CC binding reagent which comprises a glycosylphosphatidyl inositol (GPI)-  
CC anchored TGF-beta1 binding protein referred to as r150 which acts as  
CC an accessory receptor of TGF-beta. The invention is used for negatively  
CC modulating TGF-beta activity, and thus for treating conditions  
CC characterised by overproduction of TGF-beta, such as cancer. Antisense  
CC molecules of the invention are used for increasing TGF-beta availability  
CC and increase graft success. The present sequence is human r150 protein  
CC encoding DNA.  
XX  
SO Sequence 5883 BP; 1696 A; 1173 C; 1241 G; 1772 T; 1 other;  
XX

Alignment Scores:  
Pred. No.: 0 Length: 5883  
Score: 7318.50 Matches: 1426  
Percent Similarity: 98.75% Conservative: 1  
Best local Similarity: 98.69% Mismatches: 1  
Query Match: 25 Indels: 17  
DB: Gaps: 1

US-10-020-095-4 (1-1428) x AAD49435 (1-5883)  
QY 1 MetGlnGlyProProLeuLeuThrAlaAlaHsIleLeuLeuCysValCysThrAlaLeu 20  
Db 113 ATGAGGGGCCACCGCTCTGACCGCGCCACCTCTCTGCTGTCACCGCCGCTG 172  
QY 21 AlaValAlaProGlyProArgPheLeuValThrAlaProGlyIleIleArgProGly 40  
Db 173 GCCGTGGCTCCCGGCTCGCTTCTGTGACAGCCCAAGATCATCAAGCCGAGGA 232  
QY 41 AsnValThrIleGlyValGlnLeuLeuGlnHsIleCysProSerGlnValThrValAla 60  
Db 233 AATGTACTATTGGGGTGGAGCTTCTGGAACACTGCCCTTCACAGGTGAGTGAAGGCG 292  
QY 61 GlnLeuLeuLysThrAlaSerAsnLeuThrValSerValLeuGlnAlaGlnValPhe 80  
Db 293 GAGCTGCTCAAGACAGCATCAACCTACTGCTCTCTGAGAGCAGAAAGATCTTT 352  
QY 81 GlnLysGlySerPheLysThrLeuThrLeuProSerLeuProLeuAsnSerAlaAspGln 100  
Db 353 GAAGAAAGCTCTTTTAAGACACTTACTCTTCATCACTACTGAAACAGTCAGATGAG 412  
QY 101 IleTyrGlnLeuLysValThrGlyArgThrGlnAspGlnIleLeuPheSerAsnSerThr 120  
Db 413 ATTATGACTAGCTGTAACCGGACGACCAAGATGAGATTATTTCTTAATATGATACC 472  
QY 121 ArgLeuSerPheGlnThrLysArgIleSerValPheIleGlnThrAspLysAlaLeuTyr 140  
Db 473 CGCTTATCATTTAGACCAAGAGAAATATCTGTTCAATTAACAGCAAGCCCTTATAC 532  
QY 141 LysProLysGlnGlnValLysPheArgIleValThrLeuPheSerAspPheLysProTyr 160  
Db 533 AAGCCAAAGCAAGAGTGAAGTTTCCGATTTACACTCTTCTCAGATTTTAAGCCTTAC 592  
QY 161 LysThrSerLeuAsnIleLeuIleLysAspProLysSerAsnLeuIleGlnIntyrPhe 180



Db 593 AAAACCTCTTTAAACATCTCATTAAGACCCCAAAATCAAAATTGATCCAAACAGTGGTTG 652  
Qy SerGlnGlnSerAspLeuGlyValIleSerLysThrPheGlnLeuSerSerHisProIle 200  
Qy 181  
Db 653 TCACAAACAAGATGATTTGGAGTCATTTCCAAAACCTTTCAGCTATCTCCCATCCAAATA 712  
Qy 201 LeuGlyAspTyrSerIleGlnValGlnValAsnAspGlnThrTyrTyrGlnSerPheGln 220  
Db 713 CTGGGAGATGGTCTATTCAGATTCAGATGAATGACACAGACATATATCATCATTCATTGAG 772  
Qy 221 ValSerGlnTyrValLeuProLysPheGlnValThrLeuGlnThrProLeuTyrCySer 240  
Db 773 GTTTCAGAAATGATATATACCAAAATTTGAAGTACCTTTCAGACACCAATTAATATGTTCT 832  
Qy 241 MetAsnSerLysHisLeuAsnGlyThrIleThrAlaLysTyrThrTyrGlyLysProVal 260  
Db 833 ATGAAATTCATAGCATTTAAATGATGACCAACGCGCAAGTATACATATGGAGAGCCAGTG 892  
Qy 261 LysGlyAspValThrLeuThrPheLeuProLeuSerPheTyrGlyLysLysAsnIle 280  
Db 893 AAAGAGACGTAAACGCTTACATTTTACCTTATCTTTGGGGAAGAAGAAAATATT 952  
Qy 281 ThrLysThrPheLysIleAsnGlySerAlaAsnPheSerPheAsnAspGlnLysLys 300  
Db 953 ACAAAACATTTAAAGATTAATGATCTGCAAAATCTCTTTTAATGATGAGAGATGAAA 1012  
Qy 301 AsnValMetAspSerSerAsnGlyLeuSerGlnTyrLeuAspLeuSerProGlyPro 320  
Db 1013 AATGTATATGATTTCTTCAATGATGATCTTTGTGATACCTGATCTATCTTCCCTGAGCA 1072  
Qy 321 ValGlnIleLeuThrThrValThrGlnSerValThrGlyIleSerArgAsnValSerThr 340  
Db 1073 GTAGAAATTTTAAACACAGTACAGAAATCAGTTACAGATTTTCAAGAAATGTAAGCACT 1132  
Qy 341 AsnValPhePheLysGlnHisAspTyrIleIleGlnPhePheAspTyrThrThrValLeu 360  
Db 1133 AATGTGTTCTTCAAGACATATATCAATCATGATTTTGTGATTAATCTATCTGCTGTTG 1192  
Qy 361 LysProSerLeuAsnPheThrAlaThrValLysValThrArgAlaAspGlyAsnGlnLeu 380  
Db 1193 AAGCCATCTCTCAATCTCAACGACGACGTGAAGTAACTGTCGATGCGCAACCAACG 1252  
Qy 381 ThrLeuGlnGlnArgArgAsnAsnValValIleThrValThrGlnArgAsnTyrThrGln 400  
Db 1253 ACTCTTGAAGAAAGAAATTAATGATGATCAATACATGACATGACAGAGAACTTATCTGAG 1312  
Qy 401 TyrTyrSerGlySerAsnSerGlyAsnGlnLysMetGlnValGlnLysIleAsnTyr 420  
Db 1313 TACTGAGCGGATCTTAACAGTGAATAATCAGAAAATGGAAGCTGTTCAAGAAAATTAATTAT 1372  
Qy 421 ThrValProGlnSerGlyThrPheLysIleGlnPheProIleLeuGlnLysSerSerGln 440  
Db 1373 ACTGTCCCCCAAGTGAATCTTTAAGATTGAATCCCAATCTGAGAGATTCACAGTGA 1432  
Qy 441 LeuGlnLeuLysAlaTyrPheLeuGlySerLysSerSerMetAlaValHisSerLeuPhe 460  
Db 1433 CTRACATTAAGGCCATATTTCTTGTGATGAATAAGTACATGCGACTTATGTCGTTT 1492  
Qy 461 LysSerProSerLysThrTyrIleGlnLeuLysThrArgAspGlnLysIleLysValGly 480  
Db 1493 AAGTCTCCCTAGTAAAGCATACATCCAACTAAACAAAGAGATGAATAATTAAGGGGGA 1552  
Qy 481 SerProPheGlnLeuValValSerGlyAsnLysArgLeuLysGlnLeuSerTyrMetVal 500  
Db 1553 TCGCCTTTTGAAGTGTGTGTAGTGGCAAAACGATTAAAGAGATTAAATATGATGATA 1612  
Qy 501 ValSerArgGlyGlnLeuValAlaValGlyLysGlnAsnSerThrMetPheSerLeuThr 520  
Db 1613 GATTCAGAGGACAGTGTGTGTGTGAGAAAACAAAATTCACAAATGTTCTCTTTAACA 1672  
Qy 521 ProGlnAsnSerTyrThrProLysValAlaValIleValTyrTyrIleGlnLysAspGly 540  
Db 1673 CCAGAAATTTTGTGACTCCAAAAGCCTGTATATGTTATATGATGATGATGAG 1732

Qy 541 GlnIleIleSerAspValLeuLysIleProValGlnLeuValPheLysAsnLysIleLys 560  
Db 1733 GAATTTATAGGATGATCTTAAATAATCTGTGACGCTTGTTTTAAATAATAGATTAAAG 1792  
Qy 561 LeuTyrTyrSerLysValLysAlaLysProSerGlnLysValSerLeuArgIleSerVal 580  
Db 1793 CTATATGGAGTAAGTGAAGTGAACCATCTGAGAAAGCTCTCTTAGATCTCTGTG 1852  
Qy 581 ThrGlnProAspSerIleValGlyIleValAlaValAspLysSerValAsnLeuMetAsn 600  
Db 1853 ACACAGCCTGATCATAGTTGGAGTTGAGCTGTGACMAAAGTGTATCTGATGAT 1912  
Qy 601 AlaSerAsnAspIleThrMetGlnLysAsnValValIleGlnLeuGlnLysThrAsnThrGly 620  
Db 1913 GCCTTAATGATATTTATCAATGAGAAATGTGTGCTCCATGAGTTGAACCTTTATACACAGA 1972  
Qy 621 TyrTyrLeuGlyMetPheMetAsnSerPheAlaValPheGlnLysCysGlyLeuTyrVal 640  
Db 1973 TATATTTAGGATGTTGATGATTTTGGAGCTTTCAGGAATGAGACTCGGGTA 2032  
Qy 641 LeuThrAspAlaAsnLeuThrLysAspTyrIleAspGlyValTyrAspAsnAlaGlyTyr 660  
Db 2033 TTGACAGATGCAAACTCAGAGAGATTATATGATGTTGATGATGACATGACAGAAATAT 2092  
Qy 661 AlaGlnArgPheMetGlnGlnLysAsnGlnGlyHisIleValAspIleHisAspPheSerLeu 680  
Db 2093 GCTGAGAGTTATGAGAGAAATATAGAGCAATATGATGATATTCATGACATTTTCTTGG 2152  
Qy 681 GlySerSerProHisValArgLysHisPheProGlnTyrTyrIleTyrLeuAspThrAsn 700  
Db 2153 GGTACAGATCCACATGTCGAAAGCATTTTCCAGAGACTTGGATTTGGTGAGACCAAC 2212  
Qy 701 MetGlyTyrArgIleTyrGlnGlnPheGlnValThrValProAspSerIleHisSerTyr 720  
Db 2213 ATGGGTTCAGGATTTTACCAAGAAATTTGAAGTAACTGATTCATATCATCATCTTGG 2272  
Qy 721 ValAlaThrGlyPheValIleSerGlnLysLeuGlyLeuGlyLeuThrThrPheProVal 740  
Db 2273 GTGGCTACTGGTTTGTATCTCTGAGACCTGGGATCTTGAATGACATCACTCAGAG 2332  
Qy 741 GlnLeuGlnAlaPheGlnProPhePheIlePheLeuAsnLeuProTyrSerValIleArg 760  
Db 2333 GAGCTCCAGACCTTCCAAACATTTTTCATTTTGTGAATCTTCCCTACTCTGTATACGA 2392  
Qy 761 GlyGlnGlnPheAlaLeuGlnIleThrIlePheAsnTyrLeuLysAspAlaThrGlnVal 780  
Db 2393 GGTGAAGAATTTGCTTGGAAATATCATATTCATATTTGAAGATGCCACTGAGGT 2452  
Qy 781 LysValIleIleGlnLysSerAspLysPheAspIleLeuMetThrSerSerGlnIleAsn 800  
Db 2453 AAGGTAAATCATTTGAAGAAAGTGAACAAATTTGATATTCATATGATCTCAATTAATAAT 2512  
Qy 801 AlaThrGlyHisGlnGlnIleThrLeuLeuValProSerGlnAspGlyAlaThrValLeuPhe 820  
Db 2513 GCCACAGGCCACAGACACCTTGTGTTCCAGTGAAGATGGGCAACGTCTTTTTT 2572  
Qy 821 ProIleArgProThrHisLeuGlyGlnIleProIleThrValThrAlaLeuSerProThr 840  
Db 2573 CCCATCAGGCCAACATCTGGAGAAATTCCTATCAGATCAAGCTCTTCAACCCACT 2632  
Qy 841 AlaSerAspAlaValThrGlnMetIleLeuValLysAlaGlnGlyIleGlnLysSerTyr 860  
Db 2633 GCTTCTGATGCTGTACCCAGATGATTTTATGAAGGCTGAAGAAATGAATAATCATAT 2692  
Qy 861 SerGlnSerIleLeuLeuAspLeuThrAspAsnArgLeuGlnSerThrLeuLysThrLeu 880  
Db 2693 TCACATCATCTTATTTAGACTGATGACATAGAGTACACAGTAACTTGAAGAACTTTG 2752  
Qy 881 SerPheSerPheProAsnThrValThrGlySerGlnArgValGlnIleThrAlaIle 900  
Db 2753 AGTTCTCATTTTCTCTTAATATACAGTGAAGTGAAGAAAGTTCAATCATCTCAATTT 2812

QY 901 GlyAspValLeuGlyProSerIleAsnGlyLeuAlaSerLeuIleArgMetProTyrGly 920  
 Db 2813 GGAGATGTTCTTGCTTCCTTCATCAATGGCTTAGCCCTCATGATTCGGATCCCTTAATGGC 2872  
 QY 921 CysGlyGluGlnAsnMetIleAsnPheIleProAsnIleTyrIleLeuAspTyrLeuThr 940  
 Db 2873 TGTGTGTAACAAGAACATGATTAATTTGCTCCAAATATTTACATTTTGATTAATTCAGCT 2932  
 QY 941 LysIleLysGlnLeuThrAspAsnLeuLysGlyLysAlaLeuSerPheMetArgGlnGly 960  
 Db 2933 AAAAAGAAACAACGACACGATTAATTTGAAAGAAAAGCTCTTCATTTATAGGCAAGCT 2992  
 QY 961 TyrGlnArgGluLeuLeuTyrGlnArgGluAspGlySerPheSerAlaPheGlyAsnTyr 980  
 Db 2993 TACCAAGAGAACTTCTCTATACAGAGGAAATGGCTCTTCAGAGCTTTGGGAATTAAT 3052  
 QY 981 AspProSerGlySerThrTyrLeuSerAlaPheValLeuArgCysPheLeuGlnAlaAsp 1000  
 Db 3053 GACCTTCTGGAGACATGATGTTGTCAGCTTTGTTTAAAGATGTTCTTGAAGCCGAT 3112  
 QY 1001 ProTyrIleAspIleAspGlnAsnValLeuHisArgThrTyrThrTyrLeuLysGlyHis 1020  
 Db 3113 CTTATCATGATTAATTTGATCAGAAATGTGTACACAGAAATACACTTGGCTTAAGAGCAT 3172  
 QY 1021 GlnLysSerAsnGlyGluPheTyrAspProGlyArgValIleHisSerGluLeuGlnGly 1040  
 Db 3173 CAGAAATCAACGGTGAATTTTGGGATCCAGAAAGATGATTCATAGAGACTTCAAGGT 3232  
 QY 1041 GlyAsnLysSerProValThrLeuThrAlaTyrIleValThrSerLeuLeuGlyTyrArg 1060  
 Db 3233 GGCATATTAAGCTCAAGTAACCTTAACAGCTTAATGTGAACCTTCTCTGGGATTAAGA 3292  
 QY 1061 LysTyrGlnProAsnIleAspValGlnLysSerIleHisPheLeuGlnSerGluPheSer 1080  
 Db 3293 AAGATTCAGCTTAACATTTGATGTGCAGAGCTTATCCATTTTGGAGCTTCGAATTCAGT 3352  
 QY 1081 ArgGlyIleSerAspAsnTyrThrLeuAlaLeuIleThrTyrAlaLeuSerSerValGly 1100  
 Db 3353 AGAGGAATTTTCAACAATTAATTAATCTAGCCCTTAATTAATGATTTGATTCAGTGGGG 3412  
 QY 1101 SerProLysAlaLysGluAlaLeuAsnMetLeuThrTyrPheAlaGlnGlnGlnGly 1120  
 Db 3413 AGTCCTTAAGCGAAGGAAGCTTTGAATATGCTGACTTGAAGAGACAAACAAGAGTGGC 3472  
 QY 1121 MetGlnPheTyrValSerSerGluSerLysLeuSerAspSerTyrGlnProArgSerLeu 1140  
 Db 3473 ATGCATTTCTGGGTGTCATCGAGAGTCCAACTTCTGACTCTGAGACCCAGCTCCCTG 3532  
 QY 1141 AspIleGluValAlaAlaTyrAlaLeuLeuSerHisPheLeuGlnPheGlnThrSerGlu 1160  
 Db 3533 GATATTTGAAGTTGACAGCTTAATGACCTGCTCAACACTTCAACATTTCAACACTTGGAG 3592  
 QY 1161 GlyIleProIleMetArgTyrPheLysSerArgGlnArgAsnSerLeuGlyPheAlaSer 1180  
 Db 3593 GGAATCCCAATTAAGAGGTGCTTAAGCGCAAGAAATACCTTGGGGTGTTCATCT 3652  
 QY 1181 ThrGlnAspThrThrValAlaLeuLysAlaLeuSerGluPheAlaAlaLeuMetAsnThr 1200  
 Db 3653 ACTAGATACCACTGAGGCTTAAGGCTCTGCTGATTTGAGAGCCCTTAATAGAACACA 3712  
 QY 1201 GlnArgThrAsnIleGlnValThrValThrGlyProSerSerProSerPro----- 1217  
 Db 3713 GAAAGGACAAATATCCAAATGATACCGGTGACGGGGCTAGCTACCAAGTCCGTAAAGTTT 3772  
 QY 1218 -----LeuAlaValAlaGlnPro 1223  
 Db 3773 CTGATTGACACACAACCGCTTACTCTTCAGACAGACAGCTTGCTGTGTCACGCCA 3832  
 QY 1224 MetAlaValAsnIleSerAlaAsnGlyPheGlyPheAlaIleCysGlnLeuAsnValVal 1243  
 Db 3833 ATGGCAGTTAATATTTCCGCAAAATGGTTTGGATTTGCTATTTTTCAGCTCAAGTGTGTA 3892  
 QY 1244 TyrAsnValLysAlaSerGlySerSerArgArgArgSerIleGlnAsnGlnGlnAla 1263

Db 3893 TATTAATGTGAAGCTTCTGGGTCTTCTAGAAAGCGAAGATCTATCCAAATCAAGAGCC 3952  
 QY 1264 PheAspLeuAspValAlaValLysGluAsnLysAspAspLeuAsnHisValAspLeuAsn 1283  
 Db 3953 TTGATTTAATGATGTTGCTGTAAAGAAATTAAGATGATCTTCAATCATGTGGATTGCAAT 4012  
 QY 1284 ValCysThrSerPheSerGlyProGlyArgSerGlyMetAlaLeuMetGluValAsnLeu 1303  
 Db 4013 GTGTGTACAAAGCTTTTGGGCCCCGGTAGAGATGGCATGGCTCTTATAGAAATTAACTTA 4072  
 QY 1304 LeuSerGlyPheMetValProSerGluAlaIleSerLeuSerGluThrValLysLysVal 1323  
 Db 4073 TTAAAGTGCCTTAATGAGCTTCAAGAACAAATTTCTGTGACGACAGACAGTGAAGAAAGTG 4132  
 QY 1324 GluTyrAspHisGlyLysLeuAsnLeuTyrLeuAspSerValAsnGluThrGlnPheCys 1343  
 Db 4133 GAATATGATCATGGAACAACTCAACCTCTATTTAGATTTCTGTAAATGAAACCCAGTTTGT 4192  
 QY 1344 ValAsnIleProAlaValAlaArgAsnPheLysValSerAsnThrGlnAspAlaSerValSer 1363  
 Db 4193 GTTAATATTCCTGCTGTGAGAAACTTTAAAGTTTCAAAATCCCAAGATGCTTCAGTGTCC 4252  
 QY 1364 IleValAspTyrTyrGluProArgArgGlnAlaValArgSerTyrAsnSerGluValLys 1383  
 Db 4253 ATAGTGATTAATCATGAGCCAGAGAGACAGCGGTGAGAGATTACAACTGTGAAGTCAG 4312  
 QY 1384 LeuSerSerCysAspLeuCysSerAspValGlnGlyCysArgProCysGluAspGlyAla 1403  
 Db 4313 CTGCTCTCTGTGACCTTTGACATGATGTCCAGGGCTGCGCTCTGTGAGATGGAGCT 4372  
 QY 1404 SerGlySerHisHisSerSerValIlePheIlePheCysPheLysLeuTyrPhe 1423  
 Db 4373 TCAGGCTCCCATCATACACTTTCAGTCATTTTATTTCTGTTTCAAGCTTCGTATTTT 4432  
 QY 1424 MetGluLeuTyrLeu 1428  
 Db 4433 ATGCACTTGGCTG 4447  
 Db 4433 ATGCACTTGGCTG 4447  
 RESULT 7  
 AAL49816 standard; cDNA; 4335 BP.  
 AAL49816;  
 AC  
 XX  
 DT 04-DEC-2002 (first entry)  
 XX  
 DE Human platelet alloantigen Govb coding sequence.  
 KW Human; platelet alloantigen; Govb; Govb; single nucleotide polymorphism;  
 KW SNP; biallelic; bleeding disorder; post-transfusion purpura;  
 KW post-transfusion platelet refractoriness; haemostatic; vaccine;  
 KW neonatal alloimmune thrombocytopenia; gene; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 FT CDS 1..4335  
 FT /tag= a  
 FT /product= "Govb"  
 FT /partial  
 FT /note= "no stop codon"  
 FT replacement(2108,A)  
 FT /tag= b  
 XX  
 FN MO200270738-A2.  
 XX  
 PD 12-SEP-2002.  
 XX  
 PF 07-MAR-2002; 2002NC-CAD00291.  
 XX  
 PR 07-MAR-2001; 2001US-273941P.  
 XX

PA (SCHU/) SCHUH A.  
 XX Schuh A, Ouwehand W;  
 XX  
 DR MPI: 2002-713460/77.  
 DR P-PSDB; AAO19373.  
 XX  
 PT New isolated oligonucleotide binding to a region of CD109 nucleic acid  
 PT having a single nucleotide polymorphism that distinguishes a Gova  
 PT and/or Gova allele, useful for treating blood disorders e.g. autoimmune  
 PT thrombocytopenia  
 XX  
 PS Claim 7, Page 35-41; 69pp; English.  
 XX  
 CC The present invention relates to a sequence capable of binding  
 CC specifically to a CD109 nucleic acid which has a single nucleotide  
 CC polymorphism that distinguishes the Gova and Gova alleles. Detection of  
 CC the Gova genotype is useful for detecting whether the subject has or is at  
 CC risk of a blood disease, disorder or abnormal physical state, such as  
 CC bleeding, or increased risk of bleeding, due to autoimmune destruction of  
 CC blood platelets, e.g., post-transfusion purpura, post-transfusion  
 CC platelet refractoriness or neonatal alloimmune thrombocytopenia. The  
 CC nucleic acid and polypeptide are useful for Gova genotyping or phenotyping  
 CC individuals. The present sequence is the Gova coding sequence.  
 XX  
 SQ Sequence 4335 BP; 1295 A; 838 C; 922 G; 1280 T; 0 other;

## Alignment Scores:

Pred. No.: 0 Length: 4335  
 Score: 7317.50 Matches: 1426  
 Percent Similarity: 98.75% Conservative: 1  
 Best Local Similarity: 98.69% Mismatches: 1  
 Query Match: 99.58% Indels: 17  
 DB: 24 Gaps: 1

US-10-020-095-4 (1-1428) x AAL49816 (1-4335)

QY 1 MetGInGlyProProLeuLeuThraAlaAlaHsLeuLeuCyValCyThrAlaAlaLeu 20  
 DB 1 AAGCAGGGGCCACCGCTCTGACCGCGGCCACCTCTGCGTGGACCGCCGCGTGG 60  
 QY 21 AlaValAlaProGlyProAlaGlyLeuValThraAlaProGlyIleIleArgProGlyGly 40  
 DB 61 GCGGTGGCTCCGCGGCTCGGTTCTGGTGACAGCCCGAGGATCATCAGCCCGAGGA 120  
 QY 41 AsnValThrIleGlyValGluLeuLeuGluHisCysProSerGlnValThrValysAla 60  
 DB 121 AATGTGACTAATGGGGTGGAGCTTCTGGAACACTGCCCTTCAAGGTGACTGGAAGCG 180  
 QY 61 GluLeuLeuValThrAlaSerAsnLeuThraAlaSerValIleuGluAlaGluValPhe 80  
 DB 181 GAGGTGCTCAAGACATCAACCTCAGTGTCTGTCTGAGAGCAAGAGAGCTTT 240  
 QY 81 GlyIyGlySerPheIySerThreuthThreuthProSerLeuProLeuAsnSerAlaaspGlu 100  
 DB 241 GAAAAAGGCTCTTTAAGACACTTACTCTTCATCACTACCTCTGAACAGTGCAGATAG 300  
 QY 101 IleIyGluLeuValThrGlyArgThrGlnAspGluIleLeuPheSerAsnSerThr 120  
 DB 301 AATTATGAGCTACGTGTAAACCGAGCTACCCAGATGAGATTTATTTCTTAATAGTACC 360  
 QY 121 ArgLeuSerPheGluThrIySArgIleSerValPheIleGlnThraSplyValLeuTyr 140  
 DB 361 CCTTATATCTTGAAGCAAGAGAAATATCTGTCTTATTCATAACAGCAAGAGCTTATAC 420  
 QY 141 LysProLySGInGluValIySPhArgIleValThreuthPheSerAspPheIySProLy 160  
 DB 421 AAGCCAAACAGAGAGTGAAGTTTGCACTTTTACACTCTCTCAGATTTTAAGCTTAC 480  
 QY 161 LysThrSerLeuAsnIleLeuIleLysAspProLySAsnLeuIleGlnGlnTrpLeu 180  
 DB 481 AAAACCTTTTAAACATCTCTCAATTAAGAGACCCCAATCAAAATTTGATTCACAGTGGTGG 540

QY 181 SerGInGlnSerAspLeuGlyValIleSerIySThrPheGlnLeuSerSerHisProIle 200  
 DB 541 TCACAAACAAGATGATCTTGAGAGTCAATTCACAACTTTTACGATCTTCCATCACAATA 600  
 QY 201 LeuGlyAspTrpSerIleGlnValGlnValAsnAspGlnThrIyTyrGlnSerPheGln 220  
 DB 601 CTGGTGACTGTGTATTAAGATTCACAGATGATGACAGACATATTAATCAATCAATTCAG 660  
 QY 221 ValSerGluTyrValLeuProLySPhGluValThreuthGlnThrProLeuTyrCysSer 240  
 DB 661 GTTTCAGAAATGATGATTCACAAATTTGAAGTCACTTTTGCAGACACATTAATTTGTTCT 720  
 QY 241 MetAsnSerIySHisLeuAsnGlyThrIleThraAlaIySArgIySProLyVal 260  
 DB 721 ATGAATCTTAAGCATTTTAAATGATGATCAACGCGAAAGTATATATGGAAGCAAGTGG 780  
 QY 261 LysGlyAspValThreuthThreuthPheLeuProLeuSerPheTrpGlyLysLysAsnIle 280  
 DB 781 AAGGAGAGCGTACAGCTTACATTTTACCTTTATCTTTTGGGGAAGAGAAATATATTT 840  
 QY 281 ThrIySThrPheIySHisLeuAsnGlySerAlaAsnPheSerPheAsnAspGluGluMetLys 300  
 DB 841 ACAAACAATTTAAGATTAATGATGATGCAACATCTCTTTAATGATGAAGAGATGA 900  
 QY 301 AsnValMetAspSerSerAsnGlyLeuSerGluTyrIleuAspLeuSerSerProGlyPro 320  
 DB 901 AATGTAAATGATTTCTTCAATGAGACTTTCTGAATACCTCGATCATCTTCCCTGGACCA 960  
 QY 321 ValGluIleLeuThrThraValThrGluSerValThrglyIleSerArgAsnValSerThr 340  
 DB 961 GTAGAAATTTTAAACACAGTACAGAAATCACTTACAGGTATTTCAAGAAATGTAAGCACT 1020  
 QY 341 AsnValIlePhePheIySHisAspTyrIleIleGluPhePheAspTyrThrThraValLeu 360  
 DB 1021 AATGTGCTTCAAGCAACATGATTAATCACTGATTAAGTTTGTGATTAATCACTGCTCTTG 1080  
 QY 361 LysProSerLeuAsnPheThraAlaThraValIySValThraAlaAspGlyAsnGlnLeu 380  
 DB 1081 AAGCATCTTCACTCACTTCAACAGCACTGTGAAGTGAATTCGTGCTGATGAGCAACAATG 1140  
 QY 381 ThrLeuGluGluArgArgAsnAsnValIleThraValThrglnAsnTyrThrGlu 400  
 DB 1141 ACTCTTGAAGAAAGAAATTAATGATGATTAACAGAGACACAGAGAAATCACTACTAG 1200  
 QY 401 TyrTrpSerGlySerAsnSerGlyAsnGlnIySmetGluAlaValGlnIySLeuTyr 420  
 DB 1201 TACTGAGCGGATCTTAACAGTGAATCAAGAAATGGAAGCTGTTCAAGAAATTAATATAT 1260  
 QY 421 ThrValProGlnSerGlyThrPheIySIIleGluPheProIleLeuGluAspSerSerGlu 440  
 DB 1261 ACTGTCCCCCAAGTGGAACTTTTAAGTGAATTCCTCAATCTCGAGAGATTCAGATGAG 1320  
 QY 441 LeuGlnLeuIySAlaTyrPheLeuGlySerIySserSerMetAlaValHisSerLeuPhe 460  
 DB 1321 CTACAGTTGAAGGCTTATTTCTGTGTAAAGTACAGTGGAGTTCATATAGCTGTGT 1380  
 QY 461 LysSerProSerIySThrIyIleGlnLeuIySThrArgAspGluAsnIleLysValGly 480  
 DB 1381 AAGTCTCTTAATGAACATCACTCACTCACTCACTCACTCACTCACTCACTCACTCACT 1440  
 QY 481 SerProPheGluLeuValIleSerGlyAsnIySArgLeuIySLeuSerIySmetVal 500  
 DB 1441 TGGCTTTTGAAGTGGTGTAGTGGCAACAAGATTTGAAGAGTTAAGCTATATGTTA 1500  
 QY 501 ValSerArgGlyGlnLeuValAlaValGlyIySglnAsnSerThrMetPheSerLeuThr 520  
 DB 1501 GTATCAAGGAGACAGTGTGTGTGAGAAACAATTAACAATGTTCTCTTTAACA 1560  
 QY 521 ProGluAsnSerTrpThrProLySAlaCysValIleValIyTyrIleGluAspAspGly 540  
 DB 1561 CCAGAAATTTCTTGACCTCCAAAGCCGTGTAAATGTGTATTAATGAAGAGATGAGTGG 1620  
 QY 541 GluIleIleSerAspValLeuIySIIleProValGlnLeuValPheIySAsnIySIIleLys 560

Db 1621 GAAATTAATGAGTTCCTAAAAATTCCTGTCAGCTGTTTAAAAAATGAAATG 1680  
Qy LeuTYrTPSerLyValIyValAGIupProSerGluLyValSerLeuArgIleSerVal 580  
Db 1681 CTATATTGGAGTAAGTAAGTAAGTAAGTAAGTAAGTAAGTAAGTAAGTAAGTA 1740  
Qy ThrGlnProAspSerIleValIleValIleValIleValIleValIleValIleVal 600  
Db 1741 ACACAGCCTGACCTCATATGTTGGGATTTAGCTGTTGACAAAGTGTGAATCTGAT 1800  
Qy AAserAsnAspIleThMetGluAsnValIleValIleValIleValIleValIleVal 620  
Db 1801 GCCCTAATGATATTACATGAAATGATGATGATGATGATGATGATGATGATGAT 1860  
Qy TYrTYrLeuGlyMetPheMetAsnSerPheAlaValPheGlnGluCysGlyLeuTrpVal 640  
Db 1861 TATTATTAGGCATGTTATGAAATCTTTTGCAGCTTTCAGAAATGTGGACTGGGTA 1920  
Qy LeuThAspAlaAsnLeuThLyAspTYrIleAspGlyValTYrAspAsnAlaGluTYr 660  
Db 1921 TTGACAGATGCAAACTCAGCAAGATTAATATGATGATGATGATGATGATGATGAT 1980  
Qy AlaGluAspGlnMetGluGluAsnGluGlyHisIleValAspIleHisAspPheSerLeu 680  
Db 1981 GCTGAGAGCTTATGAGAAATGAAAGACATATTGATGATGATGATGATGATGATGAT 2040  
Qy GlySerSerProHisValArgIleHisAspProGluTrpIleTrpLeuAspThAsn 700  
Db 2041 GGTAGCATGTCACATGTCGAAAGCATTTTCAGAGACTTGATGATGATGATGATGAT 2100  
Qy MetGlyTYrArgIleTYrGlnGluPheGluValThValProAspSerIleThSerTrp 720  
Db 2101 ATGGGTCACAGATTTACCAAGAAATTTGAAGTAATGATGATGATGATGATGATGAT 2160  
Qy ValAlaTrpGlyPheValIleSerGluAspLeuGlyLeuGlyLeuThTrpProVal 740  
Db 2161 GTGGTACTGCTTTGTGATCTCTGAGGACCTGGCTCTGACATCAACACTACCTCCAG 2220  
Qy GluLeuGlnAlaPheGlnProPhePheIlePheLeuAsnLeuProTYrSerValIleArg 760  
Db 2221 GAGCTCCAAAGCTTCACACATTTTTCATTTTTCATTTTTCATTTTTCATTTTTCAT 2280  
Qy GlyGluGluPheAlaLeuGluIleThTrIlePheAsnTYrLeuLyAspAlaThGluVal 780  
Db 2281 GGTGAGAAATTTGCTTGGAAATACTATATTCAATATTGAAAGATGCCAGAGTT 2340  
Qy LysValIleIleGluLySerAspLyPheAspIleLeuMetThSerSerGlnIleAsn 800  
Db 2341 AAGTAATCATTTGAGAAAGTACAAATTTGATATTCTAATGACTTCAATGAATGAAT 2400  
Qy AlaThrGlyHisGlnGlnThrLeuValProSerGluAspGlyAlaThValLeuPhe 820  
Db 2401 GCCACAGGCCACAGACACCTTCTGATGCTCCAGTGGATGGGCAACTGTTCTTTT 2460  
Qy ProIleArgProThHisLeuGlyGluIleProIleThValThAlaLeuSerProThr 840  
Db 2461 CCATACAGGCCACACATCTGGAGAAATTCCTATCAAGTACAGCTTTTCCACCACT 2520  
Qy AlaSerAspAlaValThGlnMetIleLeuValIyValIleGluLySerTYr 860  
Db 2521 GCTTCTGATGCTGTCACCAAGATGATTTTAAAGGCTGAAGAAATGAAAAATCATAT 2580  
Qy SerGlnSerIleLeuLeuAspLeuThAspAsnArgLeuGlnSerThrLeuLyThLeu 880  
Db 2581 TCACATATCATATTATGACTGATGACATATAGCTACAGAGTACCTGAAAACTTTG 2640  
Qy SerPheSerPheProAsnThValThArgIleSerGluArgValGlnIleThAlaIle 900  
Db 2641 AGTTTCATATTCTCTCAATACAGTACCTGGCAGTGAAGATTCAGATACCTGCAATT 2700  
Qy GlyAspValLeuGlyProSerIleAsnGlyLeuAlaSerLeuIleArgMetProTYrGly 920

Db 2701 GGAAGTTCCTTGCTCTTCATCATGATGCTTACCTCATGATTCGGATCCTTATGCC 2760  
Qy CysGlyGluGlnAsnMetIleAsnPheAlaProAsnIleTYrIleLeuAspTYrLeuThr 940  
Db 2761 TGTGTGAACAGAAACATGAAATTTTGTGCTCCAAATATTATTAATTTGGATTATCTGACT 2820  
Qy LysIleLyGlnIleuThAspAsnLeuLyGlnIyValIleLeuSerPheMetArgGlnGly 960  
Db 2821 AAAAAGAAACATGTCAGATTAATTGAAAGAAAGAAAGCTTTTCATTTATGAGGCAAGGT 2880  
Qy TYrGlnArgGlyLeuLeuTYrGlnArgGluAspGlySerPheSerAlaPheGlyAsnTYr 980  
Db 2881 TACCAGAGAAACTTCTTATCAGAGGAAAGAGGCTTTTCAGTGTGCTTTGGAAATTA 2940  
Qy AspProSerGlySerThrTrpLeuSerAlaPheValLeuArgCysPheLeuGluAlaAsp 1000  
Db 2941 GACCCCTTGGAGACACTTGTTGTGACGCTTTGTTTAAAGATTTTCCCTTGAAGCGAT 3000  
Qy ProTYrIleAspIleAspGlnAsnValLeuHisArgThrTYrThrTrpLeuLyGlyHis 1020  
Db 3001 CCTTACATAGATATTGATCAGAAATGTTTACAGAAACATACACTTGGCTTAAAGACAT 3060  
Qy GlnLySerAsnGlyGluPheThAspProGlyArgValIleHisSerGluLeuGlnGly 1040  
Db 3061 CAGAAATCCACGGGTGAATTTGGGATCCAGAGAGTGAATTCATATGTAGCTTCAAGGT 3120  
Qy GlyAsnLySerProValThrLeuThAlaTYrIleValThSerLeuLeuGlyTYrArg 1060  
Db 3121 GGCAATTAAGTCCAGTAACACTTACAGCTTATATTTATATCTTCTCCGGGATATAGA 3180  
Qy LysTYrGlnProAsnIleAspValGlnGlySerIleHisPheLeuGlnSerGluPheSer 1080  
Db 3181 AAGATCAGGCTTAACATTGATGTGCAAGAGTATACCAATTTTGGAGTGTGAATTCAGT 3240  
Qy ArgGlyIleSerAspAsnTYrThrLeuAlaLeuIleThTYrTrIleLeuSerSerValGly 1100  
Db 3241 AGAGAAATTTCAACAATTAATCTTACCTTATTAATCTTATGATGATGATGATGATGAT 3300  
Qy SerProLyAlaLyGlnAlaLeuAsnMetLeuThTrpArgAlaGluGlnGluGly 1120  
Db 3301 AGTCCTTAAGCGAAGAAAGCTTTGAATATGCTGACTTGGAGACAAAGAAAGTGGC 3360  
Qy MetGlnPheTrpValSerSerGlnSerLyLeuSerAspSerTrpGlnProArgSerLeu 1140  
Db 3361 ATCAATTCCTGGGTGCATCAGAGTCCAAACTTTCGACTTCCTGGAGCCACTCCCTG 3420  
Qy AspIleGluValAlaAlaTYrAlaLeuLeuSerHisPheLeuGlnPheGlnThrSerGlu 1160  
Db 3421 GATATTGAAGTTCACGCTATGACAGTCTCTCACACTTCTTACAAATTTACAGACTTGTAG 3480  
Qy GlyIleProIleMetArgTrpLeuSerArgGlnArgAsnSerLeuGlyGlyPheAlaSer 1180  
Db 3481 GGAATCCCAATTTATGAGTGGCTTAAGCGCAAGAAATGACTTGGGTGGTTTGCATCT 3540  
Qy ThrGlnAspThrThValAlaLeuLyValLeuSerGlnPheAlaIleLeuMetAsnThr 1200  
Db 3541 ACTCAGAGTACCACTGGCTTTAAAGGCTCTGTGAAATTTTACAGCCCTTAATGATAC 3600  
Qy GlnArgThrAsnIleGlnValThValThArgIleProSerSerProSerPro----- 1217  
Db 3601 GAAAGGACAAATATCCAAAGTGAAGCGGCGCTTACCAAGTCTCTGTAAGATT 3660  
Qy -----LeuAlaValIleGlnPro 1223  
Db 3661 CTGATGACACACACACCGCTTACTCTTACAGACGACAGCTTGTGCTGATACAGCCA 3720  
Qy MetAlaValAsnIleSerAlaAsnGlyPheGlyPheAlaIleCysGlnLeuAsnValVal 1243  
Db 3721 ATGGCAGTTAATATTCCGGAATATGTTTGGATTGTGATTTGATTTGACGCTCAATGTGTA 3780  
Qy TYrAsnValIyValAspGlySerSerArgArgArgSerIleGlnAsnGlnGluAla 1263  
Db 3781 TATTAATGTGAAGCTTCTGGGTCTTCTAGAAAGCAAGATCTATCCAAATCAAGAGCC 3840

QY 1264 PheAspLeuAspValAlaValIysGluAsnIysAspAspLeuAsnHisValAspLeuAsn 1283  
 DB 3841 TTTGATTTAGATGCTGCTGTAATAAGAAATAAGATGATCATCATGTGATTTGAAAT 3900  
 QY 1284 ValCysThrSerPheSerGlyProGlyYargSerGlyMetAlaLeuMetGluValAsnLeu 1303  
 DB 3901 GTGTAGACAGCTTTTCGGGCCCGGGTAGAGTGCATGCTCTTAATGAAATTAACTTA 3960  
 QY 1304 LeuSerGlyPheMetValProSerGlyAlaIleSerLeuSerGlyThrValIysIysVal 1323  
 DB 3961 TTAAGTGGCTTTATGTGTGCTTCAAGACAAATTTCTTACGACAGACAGTAAAGAAAGTG 4020  
 QY 1324 GluIYrAspHisGlyIysLeuAsnLeuIYrLeuAspSerValAsnGluThrGlnPheCys 1343  
 DB 4021 GAATATGATCATGAGAAACTCAACCTCTATTAGATTCTGTAATGAAACCCAGTTTGT 4080  
 QY 1344 ValAsnIleProAlaValArgAsnIleValSerAsnThrGlnAspAlaSerValSer 1363  
 DB 4081 GTTAATATTCCTGCTGTGAGAACTTTAAAGTTTCAATACCAAGATCCTTCAGTGTCC 4140  
 QY 1364 IleValAspIYrIYrGluProArgArgGlnAlaValArgSerIYrAsnSerGluValIys 1383  
 DB 4141 ATAGTGGATTAATGATGACCAAGAGACAGCGGAGTGAAGTTACACTCTGAAGTGAAG 4200  
 QY 1384 LeuSerSerCysAspLeuCysSerAspValGlnGlyCysArgProCysGluAspGlyAla 1403  
 DB 4201 CTGTCTCTCTGTCGACCTTTGCAGTATGTCAGGGGCTGCGTCTTGTGAGATGAGCT 4260  
 QY 1404 SerGlySerHisHisSerSerValIlePheIlePheCysPheIysLeuLeuIYrPhe 1423  
 DB 4261 TCAGGCTCCCATCATCATCTTCAGTCATTTTATTTCTGTTTCAAGCTCTGACTTT 4320  
 QY 1424 MetGluLeuIYrLeu 1428  
 DB 4321 ATGGAACCTTGGCTG 4335  
 RESULT 8  
 ABQ79965 ID ABQ79965 strand: cDNA; 4761 BP.  
 AC ABQ79965;  
 XX 23-DEC-2002 (first entry)  
 XX  
 DE Human CD109 K1 variant protein encoding cDNA.  
 XX  
 KW CD109; alpha2 macroglobulin; thioester; cerebroprotective; cardiant;  
 KW immunosuppressive; haemostatic; anticoagulant; thrombolytic; human;  
 KW cardiovascular; vasotropic; gene therapy; CD109 K1; variant; gene; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 XX Key Location/Qualifiers  
 FH CDS 113..4450  
 FT FT tag= a  
 FT FT /product= "CD109 K1 variant"  
 XX  
 FN WO200270696-A2.  
 PD 12-SEP-2002.  
 XX  
 PF 07-MAR-2002; 2002WO-CA00292.  
 XX  
 PR 07-MAR-2001; 2001US-273814P.  
 XX  
 PA (SCHU/) SCHUH A.  
 PA (SUTH/) SUTHERLAND R D.  
 XX  
 PI Schuh A, Sutherland RD;  
 XX  
 DR WPI; 2002-713450/77.  
 DR P-PSDB; ABB82166.

XX  
 PT New CD109 nucleic acids and polypeptides, useful in gene therapy,  
 PT particularly for treating strokes, myocardial infarctions, thrombosis,  
 PT thrombocytopenia, autoimmune diseases, or organ or bone marrow  
 PT transplantation -  
 XX  
 PS Claim 1; Fig 1b; 156pp; English.  
 XX  
 CC The invention relates to isolated nucleic acid molecules encoding CD109  
 CC polypeptides. These nucleic acid molecules include the human cDNA  
 CC sequences comprising CD109 K1, CD109 K1-H7, CD109 K15 or their variants.  
 CC CD109 is a novel member of the alpha2 macroglobulin (alpha2M)/C3, C4, C5  
 CC family of thioester-containing proteins. The CD109 polypeptides can be  
 CC expressed by standard recombinant methodology. The CD109 nucleic acid,  
 CC CD109 polypeptide or its mimetic, or the anti-CD109 antibody is useful  
 CC for treating or detecting a disease or disorder, e.g. conditions  
 CC associated with endothelial activation, platelet activation, activation  
 CC of the coagulation or fibrinolytic systems, activation of T lymphocytes  
 CC and of the complement system, quantitative or qualitative abnormalities  
 CC of platelet function, increased or impaired platelet aggregation and  
 CC activation, increased or impaired activation of the coagulation and/or  
 CC fibrinolytic systems, or impaired or increased immune activation. These  
 CC are also useful for treating cardiovascular disorders, stroke, myocardial  
 CC infarction, thrombosis, embolism, peripheral vascular disease,  
 CC thrombocytopenia, thrombocytopenia, autoimmune diseases, organ  
 CC transplantation, or bone marrow transplantation. The CD109 nucleic acids  
 CC or their antisense nucleotide sequence are useful in gene therapy for  
 CC treating these conditions. The present sequence represents the human  
 CC CD109 K1 variant cDNA sequence.  
 XX  
 SQ Sequence 4761 BP; 1419 A; 913 C; 1016 G; 1413 T; 0 other;  
 XX  
 Alignment Scores:  
 Pred. No.: 0 Length: 4761  
 Score: 7317.50 Matches: 1426  
 Percent Similarity: 98.75% Conservative: 1  
 Best Local Similarity: 98.69% Mismatches: 1  
 Query Match: 99.58% Indels: 17  
 DB: 24 Gaps: 1  
 US-10-020-095-4 (1-1428) x ABQ79965 (1-4761)  
 QY 1 MetGlnGlyProProLeuLeuThrAlaAlaHisLeuLeuCysValCysThrAlaAlaLeu 20  
 DB 113 ATGCAAGGCGCCACCGCTCTGACCGCGCCCACTCTCTGCTGTGACCGCGCGCTG 172  
 QY 21 AlaValAlaProGlyProArgPheLeuValThrAlaProGlyIleIleArgProGly 40  
 DB 173 GCCGTGGCTCCCGGGCTCGGTTTCTGTGTGACAGCCCAAGGATCATCAGCCCGGAGCA 232  
 QY 41 AsnValIThrIleGlyValGluLeuLeuGlnHisCysProSerGlnValThrValIysAla 60  
 DB 233 AATGTACTATTTGGGTGGAGCTTCTGGAACAATGCCCTTCAAGTACGTGAGGCG 292  
 QY 61 GluLeuLeuIYrThrAlaSerAsnLeuThrValSerValLeuGluAlaGluIYrValPhe 80  
 DB 293 GAGCTGCTCAAGACAGCATCAACTCTCTGTCTGTCTGAGAGCAAGAAAGACTCTTT 352  
 QY 81 GluIYrGlySerPheIYrThrLeuThrLeuProSerLeuProLeuAsnSerAlaAspGlu 100  
 DB 353 GAAAAAGGCTCTTTTAAGACACTTACTCTTCATCATCACTCGAACAAGTGCAGATGAG 412  
 QY 101 IleIYrGluLeuArgValIThrGlyArgThrGlnAspGluIleLeuPheSerAsnSerThr 120  
 DB 413 ATTTATGAGCTACGTTGTAACCGAGCTACCAAGATGAGATTTATTTCTTAATATGATAC 472  
 QY 121 ArgLeuSerPheGluThrIYrArgIleSerValPheIleGlnIYrAspIYrAlaLeuIYr 140  
 DB 473 CGCTTATCATTTGAGACCAAGATATCTGTCTTCAATCAACAAGACAGACAGCTTATAC 532  
 QY 141 LysProIYrGlnGluValIYrPheArgIleValIThrLeuPheSerAspPheIYrProIYr 160  
 DB 533 AAGCAAGCAAGAGTGAAGTTTGCATTTGTTACACTCTCTCAGATTTTAAGCTTAC 592

QY 161 LysThrSerLeuAsnIleLeuIleLysAspProLysSerAsnLeuIleGlnGlnTrpLeu 180  
 Db 593 AAAACCTCTTTAAACATTCCTCATTAAGAGACCCCAATCAAAATTGATCCAAACAGGTGG 652  
 QY 181 SerGlnGlnSerAspLeuGlyValIleSerLysThrPheGlnLeuSerSerHisProIle 200  
 Db 653 TCACAAACAAAGTGAAGTCTGGAGTCAATTCCAAACTTTTCAGCTAATCTTCCATCCACATA 712  
 QY 201 LeuGlyAspTrpSerIleGlnValGlnValAsnAspGlnThrLysTrpLysSerPheGln 220  
 Db 713 CTGGTGCCTGCTCATTCACAGTTCAAGTGAATGACGACACATATTATCAATCATTTTCAG 772  
 QY 221 ValSerGlyLysValLeuProLysPheGlnValThrLeuGlnThrProLeuLysCysSer 240  
 Db 773 GTTTCAGAAATATGATATCCAAAATTTGAAGTCACTTTCAGACACCATTTATTTGTTCT 832  
 QY 241 MetAsnSerLysBHLysLeuAsnGlyThrIleThrAlaLysLysThrLysGlyLysProVal 260  
 Db 833 ATGAATTCCTAAGCATTTAAATGCTACCATCAGCGCAAGATATACATATGGAGACCAAGTG 892  
 QY 261 LysGlyAspValIleThrLeuThrPheLeuProLysPheTrpGlyLysLysLysAsnIle 280  
 Db 893 AAAGAGACGTAAAGCTTACATTTTACCTTTATCTTTGGGGAAGAGAAAATATTT 952  
 QY 281 ThrLysThrPheLysIleAsnGlySerAlaAsnPheSerPheAsnAspGluGluMetLys 300  
 Db 953 ACAAACAACTTTAAGATTAATGAGATCTGCAAACTTCTTTTATGATGAAGAGATGAAA 1012  
 QY 301 AsnValMetAspSerSerAsnGlyLysSerGlyLysLeuAspLeuSerSerProGlyPro 320  
 Db 1013 AATGAATAGGATTTCTTCAATGAGACTTCTGAATCTCGAGATCTAATCTTCCCTGAGCA 1072  
 QY 321 ValGluIleLeuThrThrValThrGlnSerValThrGlyLysSerArgAsnValSerThr 340  
 Db 1073 GTAGAAATTTTAAACACAGTACAGAAATCAGTACAGGATATTTCAAGAAAATGAAGCACT 1132  
 QY 341 AsnValPhePheLysGlnHisAspLysIleIleGlnPheAspLysThrThrValLeu 360  
 Db 1133 AATGCTTCTTCAAGCAACATGATTAATCATTTGAGTTTGTGATTAATACAGTCTCTG 1192  
 QY 361 LysProSerLeuAsnPheThrAlaThrValLysValThrArgAlaAspGlyAsnGlnLeu 380  
 Db 1193 AAGCATCTCTCAACTTCACAGCCACTGTGAAGGAATCTCGTGAATGAGCAACCACTG 1252  
 QY 381 ThrLeuGlnLysArgAsnAsnValIleThrValThrGlnArgAsnThrThrGlu 400  
 Db 1253 ACTCTTGAAGAAAGAAATTAATGATGATCAATCAATGACACAGAGAACTAATCTGAG 1312  
 QY 401 TyrTrpSerGlySerAsnSerGlyAsnGlnLysMetGluAlaValGlnLysIleAsnTyr 420  
 Db 1313 TACTGAGCGGATCTTAACAGTGGAAATCAGAAAATGGAAGCTGTTCAAGAAAATTAATAT 1372  
 QY 421 ThrValProGlnSerGlyThrPheLysIleGlnPheProIleLeuGlnAspSerSerGlu 440  
 Db 1373 ACTGCCCCCAAGGAACTTTTAAAGATGAATCTCCAACTCTGGAGAGATTTCCACGTAG 1432  
 QY 441 LeuGlnLeuLysAlaLysPheLeuGlySerLysSerSerMetAlaValHisSerLeuPhe 460  
 Db 1433 CTACAGTTGAAGGCTATTTCTTGGTGAAGTAAGTACATGACATGCAATCAATAGTCTGTT 1492  
 QY 461 LysSerProSerLysThrLysIleGlnLeuLysThrArgAspGluAsnIleLysValGly 480  
 Db 1493 AAGTCTCTTAAGTAAACATCAATCACTAAATAAACAAGATGAATAATTAAGTGGGA 1552  
 QY 481 SerProPheGlnLeuValValSerGlyAsnLysArgLeuLysGlnLeuSerLysMetVal 500  
 Db 1553 TCGCCTTTTGAAGTGGTGTAGTGGCAAAACGATTTGAAGAGGTAAAGTATATAGTGA 1612  
 QY 501 ValSerArgGlyGlnLeuValAlaValGlyLysGlnAsnSerThrMetPheSerLeuThr 520  
 Db 1613 GTATCAGGGGACAGTGGTGGCTGTAGAAAAAATAATTCACAAATGTTCTTTTAAACA 1672

QY 521 ProGluAsnSerTrpThrProLysAlaCysValIleValLysTyrLysIleGluAspAspGly 540  
 Db 1673 CCAAGAAATTTCTGAGCTCCAAAAGCCTGTGATTTGTGATTAATTTGAAGATGATGGG 1732  
 QY 541 GluIleIleSerAspValIleLysValIleProValGlnLeuValPheLysAsnLysIleLys 560  
 Db 1733 GAAATTTAAAGTATGTTCTTAAATAATTCCTGTCAGCTGTGTTTAAATAAAGATTAAG 1792  
 QY 561 LeuTyrTrpSerLysValLysAlaGluProSerGluLysValSerLeuArgIleSerVal 580  
 Db 1793 CTATATTGAGATTAAGTAAGTAAGTGAACCATCTGAGAAATCTCTCTAGAGATCTCTGTG 1852  
 QY 581 ThrGlnProAspSerIleValGlyIleValAlaValAspLysSerValAsnLeuMetAsn 600  
 Db 1853 ACACAGCCTGATCCATAGTGGATTTGTAGCTGTGACAAAGTGTGAATCTGATGAAT 1912  
 QY 601 AlaSerAsnAspIleThrMetGluAsnValValHisGlnLeuGlnLeuLysTrpAsnThrGly 620  
 Db 1913 GCCTTAATATATATTAACAAATGAATGTAATGTCTCAGAGTTGGAACCTTAATAACAGGA 1972  
 QY 621 TyrTyrLeuGlyMetPheMetAsnSerPheAlaValPheGlnGluCysGlyLeuTrpVal 640  
 Db 1973 TATTTATTAGGATGTTTCAATGAATTTCTTTGCACTCTTTCAGGAATGTGACTGTGGTA 2032  
 QY 641 LeuThrAspAlaAsnLeuThrLysAspTyrIleAspGlyValLysAspAsnAlaGluTyr 660  
 Db 2033 TTGACAGATCCAAACCTCAGAAAGATTAATATGATGTTTATATGACAAATGACAAATAT 2092  
 QY 661 AlaGluArgPheMetGluGluAsnGlnGlyHisIleValAspIleHisAspPheSerLeu 680  
 Db 2093 GCTGAGAGGTTTATGAGGAAATGAAGACATATTTGTAAATTTATATGATCTTTCTTTG 2152  
 QY 681 GlySerSerProHisValArgLysHisPheProGluThrTrpIleTrpLeuAspThrAsn 700  
 Db 2153 GGTAGCATCCACATGTCCGAAGACATTTTCCAGAGATCTTGATTTGCTAGACACCAAC 2212  
 QY 701 MetGlyTyrArgIleTyrGlnGlnPheGlnValThrValProAspSerIleThrSerTyr 720  
 Db 2213 ATGGGTTCCAGATTTTAAACAAATTTGAAGTAACTGATCTGATTTCACTCATCTTCTGG 2272  
 QY 721 ValAlaThrGlyPheValIleSerGluAspLeuGlyLysLeuThrThrThrProVal 740  
 Db 2273 GTGGCTACTGGTTTGTGATCTCGAGACCTGGGCTTTGGACTTAACATCTCTCCAGTG 2332  
 QY 741 GluLeuGlnAlaPheGlnProPhePheIlePheAsnTyrLeuLysAspAlaThrGluVal 760  
 Db 2333 GAGCTCCAGGCTTCCAAACATTTTTCATTTTTCATTTTTCCTTCTGCTGTATCACGA 2392  
 QY 761 GlyGluGluPheAlaLeuGlnIleThrIlePheAsnTyrLeuLysAspAlaThrGluVal 780  
 Db 2393 GGTGAAGAAATTTCTTGGAAATTAACATATTAATTAATTTGAAGTGCACATGAGTT 2452  
 QY 781 LysValIleIleGlnLysSerAspLysPheAspIleLeuMetThrSerSerGluIleAsn 800  
 Db 2453 AAGTATCAATTAAGAAATGACAAATTTGATTAATCTAATGACTTCAAAATGAATAAAT 2512  
 QY 801 AlaThrGlyHisGlnGlnThrLeuLeuValProSerGluAspGlyAlaThrValLeuPhe 820  
 Db 2513 GCCACGGCCACACAGCAGACCTTCTGTTCCAGTGAAGATGGGCACTGTTCTTTT 2572  
 QY 821 ProIleArgProThrHisLeuGlyGluIleProIleThrValThrAlaLeuSerProThr 840  
 Db 2573 CCCATCAGGCCAACACATCTGGGAAATTCCTATCAAGTCAAGCTCTTTCACCCACT 2632  
 QY 841 AlaSerAspAlaValThrGlnMetIleLeuValLysAlaGluGlyIleGlnLysSerTyr 860  
 Db 2633 GCTTCTGATCTCTCCACAGATGATTTTAAAGCTGTAAGAAATGAAGAAATATCATAT 2692  
 QY 861 SerGlnSerIleLeuLeuAspLeuThrAspAsnArgLeuGlnSerThrLysLysThrLeu 880  
 Db 2693 TCACAAATCAATTTATTAAGCTTGAATGACATGAATAGGCTACAGATCTTCAAAACCTTGG 2752  
 QY 881 SerPheSerPheProAsnThrValThrGlySerGluArgValGlnIleThrAlaIle 900



Db	2753	AGTTCTCATTTCCCTCTAATACAGTGA	CTGGCAGTGA	AAAGTTTCAGATCATCGCAATT	2812
Qy	901	GIYAspValLeuGlyProSerIleasnGlyLeuAlaSerLeuIleargMetProTyrGly			920
Db	2813	GGAAATGCTTCCTGGATCCCTTCATCAATGGCTTACCTCATGATTGGATCCCTTAATGGC			2872
Qy	921	CysGlyGluGlnAsnMetIleasnPheAlaProAsnIleTyrIleLeuAspTyrLeuThr			940
Db	2873	TGTGTGTGAACGAACATGATATAATTTTGGCTCCAAATATATTAACATTTTGGATTATCTGACT			2932
Qy	941	LysIshlySGlnLeuThrAspAsnLeuGlyGlnValaLeuSerPheMetArgGlnGly			960
Db	2933	AAAAAGAAACAACTGACGAGATTAATTGAAAGAAAAGCTCTTTCATTTATAGGCAAGCT			2992
Qy	961	TyrGlnArgGluLeuLeuTyrGlnArgGlnAspGlySerPheSerAlaPheGlyAsnTyr			980
Db	2993	TACCAAGAGAGAACTTCTCTATCAGAGGAAATGGCTCTTCACAGTCTTTGGAAATAT			3052
Qy	981	AspProSerGlySerThrThrTrpLeuSerIlePheValLeuAlaGlyCysPheLeuGlnAlaAsp			1000
Db	3053	GACCTTCCTGGAGACACTTGATGTGTCAGCTTTGTATTGAATGTTCCTTGAAGCCGAT			3112
Qy	1001	ProTyrIleAspIleAspGlnAsnValLeuIshargThrTyrThrTrpLeuLysGlyHis			1020
Db	3113	CCTTACATAGATATTGATCAGAAATGTGTTACACAGAACATACACTGGCTTAAAGACAT			3172
Qy	1021	GlnLysSerAsnGlyGluPheTrpAspProGlyArgValIleIshSerGluLeuGlnGly			1040
Db	3173	CAGAAATCCACCGGTGAATTTTGGGATCCAGAAAGATGATTCATVATGAGACTTCAAGGT			3232
Qy	1041	GlyAsnLysSerProValThrLeuThrAlaTyrIleValThrSerLeuLeuGlyTyrArg			1060
Db	3233	GGCAATATAAAGTCCAGTAACCTTAACAGCCATATTTGTAATCTTCTCCCTGGGATATAGA			3292
Qy	1061	LysTyrGlnProAsnIleAspValGlnGluSerIleHisPheLeuGluSerGluPheSer			1080
Db	3293	AAGTATCAGCCCTTAACATTGATGTGAGAGTCATTCATTTTGGAGCTCGAATTCACT			3352
Qy	1081	ArgGlyIleSerAspAsnTyrThrLeuAlaLeuIleThrTyrAlaLeuSerSerValGly			1100
Db	3353	AGAGAAATTTGAGCAANTTAACCTTAACCTTATTAACCTTATGCAATTCATCACTGGGG			3412
Qy	1101	SerProLysAlaLysGlnAlaLeuAsnMetLeuThrTrpArgAlaGluGlnGlyGly			1120
Db	3413	AGTCTTAAGCGAAGAAAGCTTTGAATATGTCTGACTTGGAGAAGCAAGAAAGATGTGC			3472
Qy	1121	MetGlnPheTrpValSerSerGluSerIshLeuSerAspSerTrpGlnProArgSerLeu			1140
Db	3473	ATGCAAATTCGGGGTGTCAACAGAGTCCAAACTTTCTGACTCTGGAGCCACGCTCCCTG			3532
Qy	1141	AspIleGluValAlaAlaTyrAlaLeuLeuSerHisPheLeuGlnPheGlnThrSerGlu			1160
Db	3533	GATATTGAAAGTTGGACGCTTAAGCACTGCTTCACACTTCTTCAATTTCAACACTTGCAG			3592
Qy	1161	GlyIleProIleMetArgTrpLeuSerArgGlnArgAsnSerLeuGlyGlyPheAlaSer			1180
Db	3593	GGAAATCCCAATTATGAGGTGTCTTAAGCAGCGAAAGAAATTAACCTTGGGGTGGTTCGACT			3652
Qy	1181	ThrGlnAspThrThrValAlaLeuLysValaLeuSerGluPheAlaAlaLeuMetAsnThr			1200
Db	3653	ACTCAGAGTACCACTGTGGCTTTAAAGGCTGTGCTGAATTTGACAGCCCTTAATGAATACA			3712
Qy	1201	GluArgThrAsnIleGlnValThrValThrGlyProSerSerProSerPro-----			1217
Db	3713	GAAAGCAAAATATCAAGTACCGTGCAGGGGCTTACCTACCAAGTCTCTGTAAAGTT			3772
Qy	1218	-----LeuAlaValaValGlnPro			1223
Db	3773	CTGATTGACACACAACCGCTTACTCTTCAGACAGACAGAGCTTGGCTGTGTACAAGCA			3832
Qy	1224	MetAlaValaAsnIleSerIleAsnGlyPheGlyPheAlaIleCysGlnLeuAsnValaVal			1243

Db	3833	ATGCACGATTAATATTCGCCGAANAGTTTGGATTTCATTTGTGACGCATAGTTGTA	3892
Oy	1244	TyrAsnValIylbAlaSerGlySerSerArgArgArgSerIleGlnaengInGluAla	1263
Db	3893	TATATATGGAAGGCTTCGGGCTCTTCAGAGAAGCAAGATCTATCCAAAATCAAGAGCC	3952
Oy	1264	PheAspLeuAspValAlaValIylbGluAsnLysAspAspLeuAsnHISValAspLeuAsn	1283
Db	3953	TTTATATTAGATGTTGCTGTAAGAAAGAAATTAACATATCTCAATCAATGTGATTTGAT	4012
Oy	1284	ValCysThrSerPheSerGlyProGlyArgSerGlyMetAlaLeuMetGluValAsnLeu	1303
Db	4013	GTGTGTACAGACTTTTGCGGCCCGGTAAGAGATGGCATGGCTTTATGGAAGTTAACTA	4072
Oy	1304	LeuSerGlyPheMetValProSerGluAlaIleSerLeuSerGluThrValLysLysVal	1323
Db	4073	TTAAGTGGCTTTATAGTGCTTCAGAGCAATTTCTCTGAGCGGAGACAGTGAAGAAAGTG	4132
Oy	1324	GluIlyrtrAspHISGlyLysLeuAsnLeuTyrLeuAspSerValaengIlyThrGlnPheCys	1343
Db	4133	GAATATATCATGTGAAGAAACATCAACCTCTATTATGATTCGTAAATGAAGCCAGTTTGT	4192
Oy	1344	ValAsnIleProAlaValaArgaenPheLysValSeranThrGlnAspAlaSerValSer	1363
Db	4193	GTTATATTTCTGCTGTGAGAAACTTTAAAGTTTCAATACCAAGATGCTTCAGTGTCC	4252
Oy	1364	IleValaAspTyrTyrGluProArgArgGlnAlaValaArgSerTyrAsnSerGluValLys	1383
Db	4253	ATAGTGATTAATCATGTAGCCAAAGGAGACAGCGGTGAGAAAGTTACACTCGAAGTGAG	4312
Oy	1384	LeuSerSerCysAspLeuCysSerAspValGlnGlyCysArgProCysGluAspGlyAla	1403
Db	4313	CTGTCCCTCCGTGACCTTTGCAGAGAGTCCAGGGCTCCGTCTGTGAGAGATGAGCT	4372
Oy	1404	SerGlySerHISHisHisSerSerValIlePheIlePheCysPheLysLeuLeuTyrPhe	1423
Db	4373	TCAGGCTCCCATCAATCACTCTTCAGTCATTTTATTTTCTGTTCAGAGCTTGTACTTT	4432
Oy	1424	MetGluLeuTyrPleu 1428	
Db	4433	ATGGAACCTTGGCTG 4447	
RESULT 9			
ABQ79967	standard; cDNA; 5895 BP.		
ID	ABQ79967		
AC	ABQ79967;		
XX	23-DEC-2002	(first entry)	
DT	Human CD109	K1-H7 variant protein encoding cDNA.	
DE	Human CD109	K1-H7 variant protein encoding cDNA.	
XX	CD109; alpha2 macroglobulin; thioester; cerebroprotective; cardiant;		
KW	Immunosuppressive; haemostatic; anticoagulant; thrombolytic; human;		
KW	cardiovascular; vasotrophic; gene therapy; CD109 K1-H7; variant; gene; ss.		
OS	Homo sapiens.		
XX			
FH	Key	Location/Qualifiers	
FT	CDS	113..4450	
FT		/*tag= a	
FT		/product= "CD109 K1-H7 variant"	
XX			
XX	W0200270696-A2.		
XX	12-SEP-2002.		
XX	07-MAR-2002; 2002W0-CA00292.		
XX	07-MAR-2001; 2001US-273814P.		
XX			
XX	(SCHU/) SCHUH A.		
XX	(SUTH/) SUTHERLAND R D.		

XX Schuh A, Sutherland RD;  
 PI WPI: 2002-713450/77.  
 DR P-PSDB: ABB82168.  
 XX  
 PT New CD109 nucleic acids and polypeptides, useful in gene therapy,  
 PT particularly for treating strokes, myocardial infarctions, thrombosis,  
 PT thrombocytopenia, autoimmune diseases, or organ or bone marrow  
 PT transplantation -  
 XX  
 PS Claim 1, Fig 2b, 156pp, English.  
 XX  
 CC The invention relates to isolated nucleic acid molecules encoding CD109  
 CC polypeptides. These nucleic acid molecules include the human cDNA  
 CC sequences comprising CD109 KI, CD109 KI-H7, CD109 K15 or their variants.  
 CC CD109 is a novel member of the alpha2 macroglobulin (alpha2u)/C3, C4, C5  
 CC family of thioester-containing proteins. The CD109 polypeptides can be  
 CC expressed by standard recombinant methodology. The CD109 nucleic acid,  
 CC CD109 polypeptide or its mimetic, or the anti-CD109 antibody is useful  
 CC for treating or detecting a disease or disorder, e.g. conditions  
 CC associated with endothelial activation, platelet activation, activation  
 CC of the coagulation or fibrinolytic systems, activation of T lymphocytes  
 CC and of the complement system, quantitative or qualitative abnormalities  
 CC of platelet function, increased or impaired platelet aggregation and/or  
 CC activation, increased or impaired activation of the coagulation and/or  
 CC fibrinolytic systems, or impaired or increased immune activation. These  
 CC are also useful for treating cardiovascular disorders, stroke, myocardial  
 CC infarction, thrombosis, embolism, peripheral vascular disease,  
 CC thrombocytopenia, thrombocytopenia, autoimmune diseases, organ  
 CC transplantation, or bone marrow transplantation. The CD109 nucleic acids  
 CC or their antisense nucleotide sequence are useful in gene therapy for  
 CC treating these conditions. The present sequence represents the human  
 CC CD109 KI-H7 variant cDNA sequence.  
 XX  
 SQ Sequence 5895 BP, 1708 A, 1174 C, 1241 G, 1772 T, 0 other;

## Alignment Scores:

Pred. No.: 0 Length: 5895  
 Score: 7317.50 Matches: 1426  
 Percent Similarity: 98.75% Conservative: 1  
 Best Local Similarity: 98.69% Mismatch: 17  
 Query Match: 99.58% Indels: 17  
 DB: 24 Gaps: 1

US-10-020-095-4 (1-1428) x ABQ79967 (1-5895)

QY 1 MetGlnGlyProProLeuLeuThraAlaHisLeuLeuCyValCyThrAlaAlaLeu 20  
 Db 113 ATGCAGGGCCACCGCTCTGACCGCGCCACCTCTCTGCTGTGACCGCGCGCTG 172  
 QY 21 AlaValAlaProGlyProArgPheLeuValThraAlaProGlyIleIleArgProGly 40  
 Db 173 GCCCGGCTCCCGGCGCTGCTGTGTGTGACAGCCCGAGGATCATCAGCCCGAGGA 232  
 QY 41 AsnValThrIleGlyValGluLeuLeuGluHisCysProSerGlnValThrValIleVal 60  
 Db 233 AATGTGACTATTGGGGTGGAGCTTCTGSAACATGCTTCTTCAAGGTACTGTAAAGCG 292  
 QY 61 GluLeuLeuLysThrAlaSerAsnLeuThraValSerValLeuGluIleGluGlyValPhe 80  
 Db 293 GAGGTGCTCAAGACGATCAAACTCACTGCTCTGTCTGTGAAGCAGAGAGGTCTTT 352  
 QY 81 GluLysGlySerPheLysThrLeuThraLeuProSerLeuProLeuAsnSerAlaAspGlu 100  
 Db 353 GAAAAAGGCTCTTTTAAACACTTACTCTTCATCTCACTTCAAGACAGTGCAGATGAG 412  
 QY 101 IleThrGluLeuArgValThrGlyValArgThrGlnAspGluIleLeuPheSerAsnSerThr 120  
 Db 413 ATTATGAGCTACGTGTAAACGAGACGTACCCAGAGTGGATTTATTTCTTAATAGTACC 472  
 QY 121 ArgLeuSerPheGluThrLysValArgIleSerValPheIleGlnThrAspLysAlaLeuTyr 140

Db 473 CGCTTATCATTTGAGACCAAGAGATAATCTGCTTCAATTCAAACAGACGCTTATAC 532  
 QY 141 LysProLysGlnGluValLysPheArgIleValThrLeuPheSerAspPheLysProTyr 160  
 Db 533 AAGCCAAAGCAAGAGTAGAGTTTCCGACTTGTACACTCTTCTTCAGATTTTAAAGCCTTAC 592  
 QY 161 LysThrSerLeuAsnIleLeuIleLysAspProLysSerAsnLeuIleGlnIntProLeu 180  
 Db 593 AAAACCTCTTAAACATTTCAATTAAGAGACCCAAATCAATTTGATCCAAACAGTGGTTG 652  
 QY 181 SerGlnGlnSerAspLeuGlyValIleSerLysThrPheGlnLeuSerSerHisProIle 200  
 Db 653 TCACCAACAAAGTATCTTGGAGTCATTTCCAAAACCTTTAGCTATCTTCCATCCATA 712  
 QY 201 LeuGlyAspTrpSerIleGlnValGlnValAsnAspGlnThrTyrTyrGlnSerPheGln 220  
 Db 713 CTGGTACTGCTGCTTATTCAGAGTTCAAGTAAATGACCAAGCATTTTATCATCATTTACAG 772  
 QY 221 ValSerGluTyrValLeuProLysPheGluValThrLeuGlnThrProLeuTyrCysSer 240  
 Db 773 GTTTCAGATATATGATATTCACAAATTTGAAGTACCTTGACAGACACCATTAATATTGCT 832  
 QY 241 MetAsnSerLysHisLeuAsnGlyThrIleThrAlaLysTyrThrTyrGlyLysProVal 260  
 Db 893 AAAGAGACGTAAACGCTTACATTTTTCCTTTTGGGAAAGAAAGAAATATTT 952  
 QY 281 ThrLysThrPheLysIleAsnGlySerAlaAsnPheSerPheAsnAspGluIleLys 300  
 Db 953 ACAAAACATTTAAGATTAATATGATCGCAACCTTCTTTATATGATGAAGATGATAA 1012  
 QY 301 AsnValMetAspSerSerAsnGlyLeuSerGluTyrLeuAspLeuSerSerProGlyPro 320  
 Db 1013 AATGTATGATTCCTTCAAAATGACCTTTCGAATACCTGTGATCTATCTTCCCTGAGCA 1072  
 QY 321 ValGlnIleLeuThrThrValThrGluSerValThrGlyIleSerArgAsnValSerThr 340  
 Db 1073 GTACAAATTTTAAACACAGTACAGCAATTCATTACAGTATTTTCAAGAAATGTACGACT 1132  
 QY 341 AsnValPhePheLysGlnHisAspTyrIleIleGluPhePheAspTyrThrThrValLeu 360  
 Db 1133 AATGTCTTCTCAAGCAACATGATTAATCATCATTTGAGTTTGTATTAATCACTGCTTG 1192  
 QY 361 LysProSerLeuAsnPheThrAlaThrValLysValThrArgAlaAspGlyAsnGlnLeu 380  
 Db 1193 AAGCATCTCTCAACTTCACAGCCACTGTGAAGTAACTGTCTGTATGCAACCAACTG 1252  
 QY 381 ThrLeuGluGluArgArgAsnAsnValIleThrValThrGlnArgAsnTyrThrGlu 400  
 Db 1253 ACTCTTGAAGAAAGAAATTAATGATGATCAATCACTGACACAGAGAACTTATCTGAG 1312  
 QY 401 TyrTrpSerGlySerAsnSerGlyAsnGlnLysMetGluAlaValGlnLysIleAsnTyr 420  
 Db 1313 TACTGAGCGGAGTCTAAACATGTGAATACAGAAATGGAACCTGTTCAGAAATTAATAT 1372  
 QY 421 ThrValProGlnSerGlyThrPheLysIleGluPheProIleLeuGluAspSerSerGlu 440  
 Db 1373 ACTGTCCCCCAAGTGGAACTTTTAAGATTGAATTCCTCGAGGAGATTCCAGAGAG 1432  
 QY 441 LeuGlnLeuValAlaTyrPheLeuGluSerLysSerSerMetAlaValHisSerLeuPhe 460  
 Db 1433 CTACAGTTGAAGGCTATTTCTTGTGATGAAGATGCAATGGAGTTCAATAGCTCTTTT 1492  
 QY 461 LysSerProSerLysThrTyrIleGlnLeuLysThrArgAspGluAsnIleLysValGly 480  
 Db 1493 AAGCTCTCTAGTAAGACATACATCCAACTAAAACAGAGATGAAAATATTAAGTGGGA 1552  
 QY 481 SerProPheGluLeuValValSerGlyAsnLysArgLeuLysGluLeuSerTyrMetVal 500  
 Db 1553 TCGCCTTTTGAAGTTGGGTAGTGCAACAAACGATTTGAAGAGATTAAGCTATATGTGA 1612

OY	501	ValSerAaGgIyGInLeuValAlaValGlyyGInaNSerThmPheSerLeuThr	520
Db	1613	GTATCCAGGGAGCAGTGTGGTGGTAGGAAAAAATAATTCACAAATGTTCTCTTTAAACA	1672
OY	521	ProGluAaNSerTrpThrProValaCySerValIleValTYrTYrIleGluAspAspGly	540
Db	1673	CCAGAAATTTCTTGACATCCAAAAACCCTGTGAATTGTATATTATTTGAAAGATGATGGG	1732
OY	541	GluIleIleSerAspValLeuIysIleProValGInLeuValPheLysAsnLysIleLys	560
Db	1733	GAATTTAATAGATGATGTTCTAAAAATTCCTGTCAGCTTGTTTTTAAATAATAGATAAG	1792
OY	561	LeuTYrTrpSerLysValLysAlaGluProSerGluLysValSerLeuArgIleSerVal	580
Db	1793	CTATTATAGATGATAAGGAAAGCTGAACATCTGAGAAAGTCCTCTTAGATCTCTGTG	1852
OY	581	ThrGInProAspSerIleValGlyIleValAlaValAspLysSerValAsnLeuMetAsn	600
Db	1853	ACACAGCTGACTCCATATGTTGGGATTTGACTGTTCGACAAAGTGTAAATCTGATGAT	1912
OY	601	AlaSerAsnAspIleThrMetGluAsnValAlaHisGluLeuGluLeuTYrAsnThrGly	620
Db	1913	GCCTCTAATGATATATTAACAATGAAAAATGTGGCCATGAGTTGGAACCTTATTAACAACA	1972
OY	621	TYrTYrLeuGlyMetPheMetAsnSerPheAlaValPheGInGluCyGlyLeuTrpVal	640
Db	1973	TATTATTATGACATGTTCATGATAATCTTTGCAGCTTCCTGAGAAATGGACCTCGGGTA	2032
OY	641	LeuThrAspAlaAsnLeuThrLysAspTYrIleAspGlyValTYrAspAsnAlaGluTYr	660
Db	2033	TTGACAGATGCAAACCTCCAGAGATTAATTGTATGGTGTTTATGACAAATCCAGATAT	2092
OY	661	AlaGluArgPheMetGluGluAsnGluLysIleValAspIleHisAspPheSerLeu	680
Db	2093	GCTGAGAGGTTTATGACAGAAAAATGAAGACATATTGATGATATTCATGACTTTCTTTG	2152
OY	681	GlySerSerProHisValArgLysHisPheProGluTrpTrpIleTrpLeuAspThrAsn	700
Db	2153	GGTAGCATCCACATGCTCCGAAGCATTTTCCAGAGACTGTGATTTGGCTGACACCAAC	2212
OY	701	MetGlyTYrArgIleTYrGInGluPheGluValThrValProAspSerIleThSerTrp	720
Db	2213	ATGGGTTCCAGATTTTACCAGAAATTTGAAGTAAGTGAACCTGATCTATACCTTTGGG	2272
OY	721	ValAlaThrGlyPheValIleSerGluAspLeuGlyLeuGlyLeuThrThrProVal	740
Db	2273	GTGGCTACTGTTTGTATCTTCGAGGAGCCTGGGCTCTTGACCTAACACTATCCACGTG	2332
OY	741	GluLeuGlnAlaPheGInProPhePheIlePheLeuAsnLeuProTYrSerValIleArg	760
Db	2333	GAGCTCCAGAGCTTCCAAACCAATTTTCTAATTTTGAATCTTCCCTACTCTGTTATCGA	2392
OY	761	GlyGInGluPheAlaLeuGluIleThrIlePheAsnTYrLeuLysAspAlaThrGluVal	780
Db	2393	GGTAGAGATTTGCTTTGGAAATACATAATTCAATTATTGAAAGATGCCACTGAGGTT	2452
OY	781	LysValIleIleGluLysSerAspLysPheAspIleLeuMetThrSerSerGluIleAsn	800
Db	2453	AAGGTAATCATTTGAGAAAGTAGACAAATTTGATATTCTTAATGACTTCAAAAGAAATAAT	2512
OY	801	AlaThrGlyHisGInGInThrLeuLeuValProSerGluAspGlyAlaThrValLeuPhe	820
Db	2513	GCACAGGCCACCAAGCAACCCCTTCTGGTCTCCAGTGAAGAGGGCAACGTGCTCTTTT	2572
OY	821	ProIleArgProThnHisLeuGlyGlyIleProIleThrValThnAlaLeuSerProThr	840
Db	2573	CCCATCAGGCCAACAACATCTGGAGAAATTTCTATCACAAGTCAACAAGCTTTTACCCACT	2632
OY	841	AlaSerAspAlaValThrGluMetIleLeuValLysAlaGluGlyIleGluLysSerTYr	860
Db	2633	GCCTCTAGTGTGTCACCCAAGATGATTTTATGAAGGCTGAAGGAATGAAAAATCATAT	2692

QY	861	GerGIInSerIleLeuLeuAspLeuThrAspAsnArgLeuGlnSerThrLeuLeu	880
Db	2693	TCACATCATCATTAATAGACTTACGACAAATGGCTACGAGATACCCTGAACACTTGG	2752
QY	881	SerPheSerPheProProAsnThrValThrGlySerGluArgValGlnIleThrAlaIle	900
Db	2753	AGTTTCATTTCTCTCTCAATACAGTGACTGGCACTGAAGAGTTGACATCAGCAATT	2812
QY	901	GlyAspValLeuGlyProSerIleAsnGlyLeuAlaSerLeuIleArgMetProTyrGly	920
Db	2813	GGAATGTTCTTGGCTCTTCATCAATGGCTTACCTCATGATTCCGATGCTTATGGC	2872
QY	921	CysGlyGluGlnAsnMetIleAsnPheAlaProAsnIleTyrIleLeuAspTyrLeuThr	940
Db	2873	TGTGTGTGAACGAACATGATTAATTTGGCTCCCAATATTTTACATTTGGATTATCTGCT	2932
QY	941	LysIleValSerGlnLeuThrAspAsnLeuLeuGlyValLeuSerPheMetArgGlnGly	960
Db	2933	AAAAAGAAACACTGACGATTAATTTGAAGAAAGAAAGCTCTTCAATTATATGAGCAAGCT	2992
QY	961	TyrGlnArgGluLeuLeuTyrGlnArgGluAspGlySerPheSerAlaPheGlyAsnTyr	980
Db	2993	TACACAGAGAACTCTCTCTATCAAGAGGAAGATGGCTCTTCACTGCTTTGGGAATTAT	3052
QY	981	AspProSerGlySerThrTyrPheLeuSerAlaPheValLeuArgCysPheLeuGluAlaAsp	1000
Db	3053	GACCTTCTGGAGACACTTGGTGTCTGACGTTGTTTAAAGATGTTTCTTGAAGCCGAT	3112
QY	1001	ProTyrIleAspIleAspGlnAsnValLeuHisArgThrTyrThrTyrPheLeuGlyHis	1020
Db	3113	CTTACATATGATTAATGATCAGAAAGTGTTACACAAACATACACTGGCTTAAAGACAT	3172
QY	1021	GlnIleSerAsnGlyGluPheTyrAspProGlyArgValIleHisSerGluLeuGlnGly	1040
Db	3173	CAGAAATCCACGGTGAATTTGGGATCCAGAGAGATGATTCATATGAGCTTCAAGCT	3232
QY	1041	GlyAsnIleSerProValThrLeuThrAlaTyrIleValThrSerLeuLeuGlyTyrArg	1060
Db	3233	GGCATATAAGTCCAGTAACACTTACAGCCATATTTGTAATCTCTCTGGGATTAAGA	3292
QY	1061	LysTyrGlnProAsnIleAspValGlnGlnSerIleHisPheLeuGlnSerGluPheSer	1080
Db	3293	AAGTATCAGCTTAACATTGATGACGACAGACTATCAATTTTGGAGTCTGAATTCAGT	3352
QY	1081	ArgGlyIleSerAspAsnTyrThrLeuAlaLeuIleThrTyrAlaLeuSerSerValGly	1100
Db	3353	AGAGGATTTCCAGCAATTATACCTCTAGCCCTTATTACTTATGATGTGATCAGTGGG	3412
QY	1101	SerProIleValIleGluAlaLeuAsnMetLeuThrTyrArgAlaGlnGlnGlyGly	1120
Db	3413	AGTCTTAAGGAGGAAGACTTTGAAATATGTGCTACTGGAGAGCAAAAGAGATGGC	3472
QY	1121	MetGlnPheTyrPheValSerSerGlnSerIleLeuSerAspSerTyrGlnProArgSerLeu	1140
Db	3473	ATGCAATTTCTGGGTGATCAGATCTCAAACTTTCTACCTCTGGCAGCCACGCTCCCTG	3532
QY	1141	AspIleGluValAlaAlaTyrAlaLeuLeuSerHisPheLeuGlnPheGlnThrSerGlu	1160
Db	3533	GATTTGATGAGTTCAGCCCTATGACCTGTCTCACACTTCTTACATTTCCACACTTCGAG	3592
QY	1161	GlyIleProIleMetArgTyrPheLeuSerArgGlnArgAsnSerLeuGlyGlyPheAlaSer	1180
Db	3593	GGATTCCAATTAATGAGGTGGCTAAGCAGGCAAGAAATAGCTTGGGTGGTTTGCATCT	3652
QY	1181	ThrGlnAspThrThrValAlaLeuIleValAlaLeuSerGluPheAlaAlaLeuMetAsnThr	1200
Db	3653	ACTCAGAGATCCACGTGGCTTTAAAGCTCTGTCTCAATTTGACGCCCTAATGAATACA	3712
QY	1201	GluArgThrAsnIleGlnValIleThrValIleThrGlyProSerSerProSerPro-----	1217
Db	3713	GAAGGACAAATATCCAAATGACCGGTACCGGGCTTACGCCAAGTCTGTAAAGATT	3772
QY	1218	-----LeuAlaValAlaGlnPro	1223

Db 3773 CTGATTGACACACACCGCTTACTCCTTCAGACAGACAGACTGTGTGTACAGCCA 3832  
Qy 1224 MetAlaValaenilSerAlaenGlyPheGlyPheAlaIleCySGlnLeuAsnValaI 1243  
Db 3833 ATGGAGTTAAATATATCCGCAAAATGTTTGATTTGCTAATTTGCACTCAATAGTTGTA 3892  
Qy 1244 TyrAsnVallybAlaSerGlySerSerArgArgArgSerIleGlnAsnGlnIuaIa 1263  
Db 3893 TATATATGGAAGGCTTCTGGGCTCTTCAGAGACCAATCTATCCAAATCAAGAGGCC 3952  
Qy 1264 PheAspLeuAspValaIaVallySGlnAsnIysAspAspLeuAsnHisValaAspLeuAsn 1283  
Db 3953 TTTGATTTAGATGTTGCTGTAAGAAATGAAGATGATCTCAATCATGTCGATTTGAT 4012  
Qy 1284 ValCySerThrSerPheSerGlyProGlyArgSerGlyMetAlaIaIaMetGlnValaAsnIleu 1303  
Db 4013 GTGTGTACAACTTTTCGGGCGCCGGGTAGAGATGCGATGCTTATGGAAGTTAACCTTA 4072  
Qy 1304 LeuSerGlyPheMetValaProSerGlyIuaIaIleSerLeuSerGlyThrVallyIyVala 1323  
Db 4073 TTAAGTGCCTTTATGCTGCTCTTCAGAACCAATTTCTCGAGCGAGACAGTGAAGAAAGTG 4132  
Qy 1324 GlnTyrAspHisGlyIyValyLeuAsnIleuTyrIleuAspSerValaAsnGlyThrGlnPheCys 1343  
Db 4133 GAATATGATCATGGAACCTCAACCTCTAATTTAGATTCGTAAATGAAGAACCCAGTTTGT 4192  
Qy 1344 ValaenileProAlaValaIArgAsnPheIyValaSerAsnThrGlnAspAlaSerValaSer 1363  
Db 4193 GTTAATATTCCTGCTGAGAGAACTTTAAAGTTTCAATACCAAGATGCTTCAGTGTCC 4252  
Qy 1364 IleValaAspTyrTyrGlnProArgArgGlnIaIaValaIArgSerTyrAsnSerGlyIuaIyAs 1383  
Db 4253 ATAGTGGATTCTATGAGCCAGAGAGACAGCGGTGAAGATTTCAACTCGAAGTGAG 4312  
Qy 1384 LeuSerSerCyAspLeuCySerSerAspValaGlnGlyCyAspArgProCySGlnAspGlyAla 1403  
Db 4313 CTGTCTCTCTGTGACCTTTGACATGATGTCCAGGGCTCCCTCTTGAGAGATGAGACT 4372  
Qy 1404 SerGlySerHisHisSerSerValIlePheIlePheCyPheIyLeuLeuTyrPhe 1423  
Db 4373 TCAGGCTCCCATCAATCACTTCCTCACTCAATTTTATTTCTGTTCAGCTTCCTGACTTT 4432  
Qy 1424 MetGlnLeuTyrLeu 1428  
Db 4433 ATGGAATTGCTGCTG 4447

RESULT 10  
AAA60199 standard, DNA; 4335 BP.  
XX  
AC AAA60199;  
XX  
DT 02-FEB-2001 (first entry)  
XX  
DE Hydrophobic domain protein cDNA HP02837 isolated from HT-1080 cells.  
XX  
KW Human, secreted protein, membrane protein; hydrophobic domain;  
KW proliferation control; differentiation induction; material transport;  
KW biophysics; signal receptor; ion channel; transporter; immunostimulant;  
KW immunosuppressant; haematopoiesis regulator; chemotactic; chemokinetic;  
KW haemostatic; thrombolytic; anti-inflammatory; tumour inhibition;  
KW autoimmune disease; Alzheimer's disease; Parkinson's disease; cancer;  
KW gene therapy; ss.  
XX  
XX Homo sapiens.  
OS  
FN MO200029448-A2.  
XX  
XX 25-MAY-2000.  
XX  
XX 17-NOV-1999; 99WO-JP06412.  
XX

PR 17-NOV-1998; 98JP-0326255.  
PR 22-DEC-1998; 98JP-0364315.  
PR 16-MAR-1999; 99JP-0069811.  
PR 27-APR-1999; 99JP-0119299.  
PR 19-MAY-1999; 99JP-0138169.  
XX  
PA (SAGA) SAGAMI CHEM RES CENT.  
PA (PROT-) PROTEGENE INC.  
PI Kato S, Kimura T;  
XX WPI; 2000-387753/33.  
DR P-PSDB; AAB12127.  
XX  
PT Proteins comprising hydrophobic regions, such as secretory and membrane  
PT proteins, useful in research and diagnostics and having various  
PT activities e.g. immunomodulatory, antiinflammatory, chemokinetic,  
PT hemostatic, thrombolytic -  
XX  
PS Claim 3; Page 244-246; 410pp; English.  
XX  
CC Secretory proteins play important roles in the proliferation control, the  
CC differentiation induction, the material transport and the biophysics of  
CC cells. Membrane proteins have important roles as signal receptors, ion  
CC channels and transporters. The present sequence is the coding sequence  
CC for a human protein which has at least one hydrophobic domain. The  
CC protein encoded by the present sequence may be a secretory or a membrane  
CC protein. The encoded protein may have cytokine and cell  
CC proliferation/differentiation activity, immune stimulating or suppressing  
CC activity, haematopoiesis activity, tissue growth activity,  
CC activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic  
CC and thrombolytic activity, anti-inflammatory activity and tumour  
CC inhibition activity. The present sequence could therefore be used for  
CC treatment of autoimmune disease, Alzheimer's disease, Parkinson's  
CC disease, and cancer via gene therapy.  
XX  
SQ Sequence 4335 BP; 1295 A; 838 C; 922 G; 1280 T; 0 other;  
XX

Alignment Scores:  
Pred. No.: 0 Length: 4335  
Score: 7307.50 Matches: 1424  
Percent Similarity: 98.62% Conservative: 1  
Best Local Similarity: 98.55% Mismatches: 3  
Query Match: 99.45% Indels: 17  
DB: 21 Gaps: 1

US-10-020-095-4 (1-1428) x AAA60199 (1-4335)  
Qy 1 MetGlnGlyProProLeuLeuThrAlaIaIleLeuCyValaCyThrAlaIaIleu 20  
Db 1 ATGAGGGGCCACCGGCTCTGACCGCGCCACCTCTGCGGTGCACCGCGGCTG 60  
Qy 21 AlaValaIaProGlyProArgPheLeuValThrAlaProGlyIleIleArgProGlyGly 40  
Db 61 GCCGTGCTCCCGGCTCCGCTTCTGTGTACAGGCCCAAGAGATCAATCAGAGCCCGAGGA 120  
Qy 41 AsnValThrIleGlyValaGlnLeuGlnHisCySerProSerGlnValThrValIyAla 60  
Db 121 AATGTGCTATTTGGGGAGGCTTCTGAAACAGTCCCTTCACAGGTGATCTGAAAGCG 180  
Qy 61 GlnLeuLeuYThrAlaSerAsnLeuThrValSerValIleuGlnIaIaGlnGlyValaPhe 80  
Db 181 GAGCTGCTCAAGACAGATCAAACTCAGTCTCTGCTGAGAGCAGAAAGAGTCTTT 240  
Qy 81 GlnIyGlySerPheIyThrIleuThrIleuProSerIleuProLeuAsnSerAlaAspGln 100  
Db 241 GAAAAAGGCTCTTTTAAAGACACTTACTCTTCATCACTACTCTGAAACAGTGCAGATGAG 300  
Qy 101 IleTyrGlnLeuArgValThrGlyArgThrGlnAspGlnIleuPheSerAsnSerThr 120  
Db 301 ATTATAGCTAGCTGTGAACCGAGCGTACCAGAGATGAGATTTATTTCTTAATAGTACC 360  
Qy 121 ArgLeuSerPheGlnThrIyAsArgIleSerValPheIleGlnThrAspIyAlaIeUtyr 140

```

Db      361  CGCTTATCATTTAGACCAAGAAATATCTGCTTCAATCAACACAGGCTTATAC 420
Qy      141  LysProLysGlnGluValLysPheArgIleValThrLeuPheSerAspPheLysProTyr 160
Db      422  AGCCCAAGCAAGAAATGAAAGTTCCGATGTTTACACTCTCTCCAGATTTTAAGCTTAC 480
Qy      161  LysThrSerLeuAsnIleLeuIleLysAspProLysSerAsnLeuIleGlnIntProLeu 180
Db      481  AAAACCTCTTAAACATTCCTCATTAAGAACCCAAATCAAAATTTGATCCACAGTGTTG 540
Qy      181  SerGlnGlnSerSerLysGlnValIleSerLysThrPheGlnLeuSerSerHisProIle 200
Db      541  TCACAAACAAGATGATCTGAGTCATTTCCAAAACCTTTTCAGCTATCTTCCCATCCAAA 600
Qy      201  LeuGluAspTrpSerIleGlnValGlnValAsnAspGlnThrTyrTyrGlnSerPheGln 220
Db      601  CTTGGTGAATGATGTATTTACCAAAATTTGAAGTACCTTTGACAGACACATTAATTTCT 660
Qy      221  ValSerGluTyrValLeuProLysPheGluValThrLeuGlnThrProLeuTyrCysSer 240
Db      661  GTTTCAGAAATATGTATTTACCAAAATTTGAAGTACCTTTGACAGACACATTAATTTCT 720
Qy      241  MetAsnSerLysHisLeuAsnGlnTyrIleThrAlaLysTyrThrTyrGluLysProVal 260
Db      721  ATGAATTCCTAAGCATTTAAATGATGACATCACCGCAAAAGATATCATATGGGAAGCCAG 780
Qy      261  LysGluAspValThrLeuThrPheLeuProLysSerPheTrpGluLysLysAsnIle 280
Db      781  AAGGAGACGTTAAGCTTACATTTTACCTTTATCTTTGGGGAAGAAAGAAATATTT 840
Qy      281  ThrLysThrPheLysIleAsnGlnSerAlaAsnPheSerPheAsnAspGlnGluMetLys 300
Db      841  ACAAATAACCTTAAAGATTAATGATGCTGCAACTTCTCTTTAATGATGAAGAGATGAA 900
Qy      301  AsnValMetAspSerSerAsnGlnLysLeuSerGluTyrLeuAspLeuSerSerProGlu 320
Db      901  AATGTATGATGATCTTCAATGATGATCTTGAATACCTGATCTATCTTCCCTGAGACA 960
Qy      321  ValGluIleLeuThrValThrGlnSerValThrGlyIleSerArgAsnValSerThr 340
Db      961  GTTAGAAATTTTAAACACAGTACAGAAATCACTTACAGGATTTCAAGAAATGTAAACACT 1020
Qy      341  AsnValPhePheLysGlnHisAspTyrIleIleGluPhePheAspTyrThrThrValLeu 360
Db      1021  AATGTCTCTTCAAGCAATGATTTACATCATTTGATTTTGTATTAATCTACTGCTTGG 1080
Qy      361  LysProSerLeuAsnPheThrAlaThrValLysValThrArgAlaAspGluAsnGlnLeu 380
Db      1081  AAGCCATCTCTCAACTTCAAGCCACTGTGAAGGTAACTGTGCTGATGCAACCAACTG 1140
Qy      381  ThrLeuGlnGluLysArgArgAsnAsnValIleThrValThrGlnArgAsnTyrThrGlu 400
Db      1141  ACTCTTGAAGAAAGAAAGAAATATGTAGTCATTAACGTGACACAGAGAAACTATATCTG 1200
Qy      401  TyrTrpSerGlnSerAsnSerGlnLysAsnGlnLysMetGluAlaValGlnLysIleAsnTyr 420
Db      1201  TACTGAGGCGGATCTTACAGTGAATTCGAAATATGAAAGCTGTTTACGAAATTAATTAAT 1260
Qy      421  ThrValProGlnSerGlnTyrPheLysIleGluPheProIleLeuGlnAspSerSerGln 440
Db      1261  ACTGTCCTCCCAAGTGAACCTTTTAAAGTGAATTCCTCAATCTCGAAGGATTCCTCATG 1320
Qy      441  LeuGlnLeuLysAlaLysPheLeuGlnLysSerLysSerMetAlaValHisSerLeuPhe 460
Db      1321  CTACAGATTAAAGGCTTAATTCCTTGGTAGTAAAGATGACATGCACTTATGTCTGTTT 1380
Qy      461  LysSerProSerLysThrTyrIleGlnLeuLysThrArgAspGlnAsnIleLysValGly 480
Db      1381  AAGTCTCCCTAGTAGACATACATCCAACTTAACAAACAGATGATTAATTAAGGTGGGA 1440
Qy      481  SerProPheGlnLeuValValSerGlnLysAsnLysArgLeuLysGluLeuSerTyrMetVal 500
Db      1441  TCGCTTTTGAAGTGGTGTAGTGCAACAAACGATTTGAAGAGTAAAGCTATATGTGTA 1500
Qy      501  ValSerArgGlnGlnLeuValAlaValGlyLysGlnAsnSerThrMetPheSerLeuThr 520
Db      1501  GTATCCAGGGGACAGTTGGTGGCTGTAGGAAACAAATTCATCAATGTCTCTTTACCA 1560
Qy      521  ProGluAsnSerTrpThrProLysAlaCysValIleValTyrTyrIleGlnAspAspGly 540
Db      1561  CCAGAAATATCTTGAGCTCCAAAGCCTGTGTAATGTGTATTAATTAAGATGATGGG 1620
Qy      541  GluIleIleSerAspValLeuLysIleProValGlnLeuValPheLysAsnLysIleLys 560
Db      1621  GAAATTAATGAAGATGTTCTAAATAATCTCTGTACGCTTGTAAAAATTAAGATAAG 1680
Qy      561  LeuTyrTrpSerLysValLysAlaGluProSerGlnLysValSerLeuArgIleSerVal 580
Db      1681  CTATTTGGAGTAAAGTGAAGTGAACCTAACCATTGGAAGATCTCTTAAGATCTCTGTTG 1740
Qy      581  ThrGlnProAspSerIleValGlyIleValAlaValAspLysSerValAsnLeuMetAsn 600
Db      1741  ACACAGCCTGACTCCATAGTTGGATTTGAGCTGTGCAAAAAGTGAATCTGATGAT 1800
Qy      601  AlaSerAsnAspIleThrMetGlnAsnValAlaHisGluLeuGlnLeuTyrAsnThrGly 620
Db      1801  GCCTTAATGATTAATTAACATGGAATAATGGTCCATGATGGCACTTTATTAACACAGA 1860
Qy      621  TyrTyrLeuGlnMetPheMetAsnSerPheAlaValPheGlnGluCysGluLeuThrVal 640
Db      1861  TATTAATTAAGGATGTTCATGTAATCTTTTGACGCTTTCAGGAATGGAGACTGGGGA 1920
Qy      641  LeuThrAspAlaAsnLeuThrLysAspTyrIleAspGlyValTyrAspAsnAlaGluTyr 660
Db      1921  TTGACAGATGCAAACTCAAGAGATTAATTAATGATGCTTTAGACAAATGACAAATAT 1980
Qy      661  AlaGluArgPheMetGlnGluAsnGlnGlyHisIleValAspIleHisAspPheSerLeu 680
Db      1981  GCTGAGAGCTTATGAGAGAAATGAAGACATATTGATGAATATTCATGACTTTCTTTG 2040
Qy      681  GlySerSerProHisValArgLysHisPheProGluThrTrpIleTrpLeuAspThrAsn 700
Db      2041  GGTAGCAGTCCACAGTCCGAAAGCAATTTCCAGAGACTGTGATTTGGCTAACAACAC 2100
Qy      701  MetGlyTyrArgIleTyrGlnGlnPheGluValThrValProAspSerIleThrSerTrp 720
Db      2101  ATGGTTCACAGATTTACCAAGAAATTTGAAGTACGTAACTGATCTTATCTCTTGG 2160
Qy      721  ValAlaThrGlyPheValIleSerGlnAspLeuGlnLysLeuTyrThrThrProVal 740
Db      2161  GTGGCTACTGGTTTGTGATCTGTAGGACCTGGGCTTTGGACTTAACAACACTCTCAAG 2220
Qy      741  GluLeuGlnAlaPheGlnProPhePheIlePheLeuAsnLeuProTyrSerValIleArg 760
Db      2221  GAGCTCCAAAGCTTCCAAACATTTTCAATTTTGAATCTTCCCTACTCTGTATTCGA 2280
Qy      761  GlyGluGlnPheAlaLeuGlnIleThrIlePheAsnTyrLeuLysAspAlaThrGluVal 780
Db      2281  GGTGAAGATTTTCTTGGAAATTAATCTATTAATTTGAAGAAATGCCACTAGGTT 2340
Qy      781  LysValIleIleGlnLysSerAspLysPheAspIleLeuMetThrSerSerGluIleAsn 800
Db      2341  AAGTATACATTTGAAGAAATGACAAATTTGATTAATTAATGACTTCAAGTGAATTAAT 2400
Qy      801  AlaThrGlyHisGlnGlnThrLeuLeuValProSerGluAspGlyAlaThrValLeuPhe 820
Db      2401  GCCACAGGCGACACAGACAGCCCTTCTGTGTTCCCATGTGAGATGGGCAACTGTTCTTT 2460
Qy      821  ProIleArgProThrHisLeuGlnGluIleProIleThrValThrAlaLeuSerProThr 840
Db      2461  CCCATCAGGCCAACACATCTGGAGAAATTTCTTAACAGTACAGCTCTTTACCCACT 2520
Qy      841  AlaSerAspAlaValThrGlnMetIleLeuValLysAlaGlnGlyIleGlnLysSerTyr 860
Db      2521  GCTTCTGATGCTATCACCCAGATGATTTTATGAAGGCTGAAGAGATTAAGAAATCATAT 2580
```

QY 861 SerGlnSerIleLeuLeuAspLeuThrAspAsnArgLeuGlnSerThrLeuLysThrLeu 880  
 Db 2581 TCACATTCATCTTATTAGACTGACGACATAGCTACAGAGTACCTGAAAACTTGG 2640  
 QY 881 SerPheSerPheProPheAsnThrValThrGlySerGluArgValGlnIleThrAlaIle 900  
 Db 2641 AGTTCTCATTTCCCTCAATACAGTGAAGTGCAGTGAAGAAAGTTCAATCATCTGCAATT 2700  
 QY 901 GlyAspValLeuGlyProSerIleAsnGlyLeuAlaSerLeuIleArgMetProTyrGly 920  
 Db 2701 GGAGATGCTTGTGGCTCTTCATCAATGCTTAGCTCATTTGATTCGGATGCTTATGCG 2760  
 QY 921 CysGlyGlnGlnAsnMetIleAsnPheAlaProAsnIleTyrIleLeuAspTyrLeuThr 940  
 Db 2761 TGTGGTAAACAGAACATGATTAATTTTGTCTCAAAATATTTACATTTTGATTTATCTGACT 2820  
 QY 941 LysIleLysGlnLeuThrAspAsnLeuLysGlyLysAlaLeuSerPheMetArgGlnGly 960  
 Db 2821 AAAAAAGAAACAACGACAGATTAATTTGAAGAAAAAGCTCTTTCATTTATGAGGCAAGGT 2880  
 QY 961 TyrGlnArgGluLeuLeuTyrGlnArgGluAspGlySerPheSerAlaPheGlyAsnTyr 980  
 Db 2881 TACAGAGAGAACTTCTATACAGGAGAGAGTGGCTTTTCAAGTCTTTTGGGAAATTAT 2940  
 QY 981 AspProSerGlySerThrTyrLeuSerLysPheValLeuArgCysPheLeuGluAlaAsp 1000  
 Db 2941 GACCTCTTGGGAGACCTTGGTGTGACCTTTTGTTTAAGATGTTTCTTGGAAAGCCGAT 3000  
 QY 1001 ProTyrIleAspIleAspGlnAsnValLeuIleSarGlyThrTyrTrpLeuLysGlyHis 1020  
 Db 3001 CCTTACATAGATATGATCAGAAATGTGTACACAGAACATACACTTGGCTTAAAGGACAT 3060  
 QY 1021 GlnLysSerArgGlnGluPheTrpAspProGlyArgValIleHisSerGluLeuGlnGly 1040  
 Db 3061 CAGAAATCAACGGGAAATTTGGGATCCAGGAAAGATGATTCATAGGACCTTCAAGGT 3120  
 QY 1041 GlyAsnLysSerProValThrLeuThrAlaTyrIleValThrSerLeuLeuGlyTyrArg 1060  
 Db 3121 GGCAATAAAGTCCAGTAACTTACAGCTTAATNTGTAATCTTCTCTCTGGGATATAGA 3180  
 QY 1061 LysTyrGlnProAsnIleAspValGlnGlnSerIleHisPheLeuGlnSerGluPheSer 1080  
 Db 3181 AAGATATCAGCCTAACATGATGTGACAGAGTCAATCCATTTTGGAGTCTGAATTCAGT 3240  
 QY 1081 ArgGlyIleSerAspAsnTyrThrLeuAlaLeuIleThrTyrAlaLeuSerSerValGly 1100  
 Db 3241 AGAGAAATTTCAACAATTAATTAATCTTACGCTTAACTTAATGCAATGTCTACAGTGGG 3300  
 QY 1101 SerProLysAlaLysGluAlaLeuAsnMetLeuThrTyrPArgAlaGlnGlnGlyGly 1120  
 Db 3301 AGTCTTAAGAGGAGGAGGAGCTTTGAATATGCTGACTTGAGACGACGAAAGAGTGGC 3360  
 QY 1121 MetGlnPheTyrPvalSerSerGlnSerLysSerAspSerTyrGlnProArgSerLeu 1140  
 Db 3361 ATGCAATTTCTGGGTGTCTACAGAGTCCAACTTTCTGACTCTGGACGCCATCCCTG 3420  
 QY 1141 AspIleGluValAlaAlaTyrAlaLeuLeuSerHisPheLeuGlnPheGlnThrSerGlu 1160  
 Db 3421 GATATGTGAAGTTCAGGCTTAATGCACTGCTTCAACCTTCAATTCACAACTTCTGAG 3480  
 QY 1161 GlyIleProIleMetArgTyrPheSerArgGlnArgAsnSerLeuGlyGlyPheAlaSer 1180  
 Db 3481 GGAATCCCAATTAATGAGGTGCTAAGCAGGCAAGAAATTAATCTTGGGTGTTTGGCATCT 3540  
 QY 1181 ThrGlnAspThrThrValAlaLeuLysAlaLeuSerGluPheAlaLeuMetAsnThr 1200  
 Db 3541 ACTCAGGATACCACTGTGGCTTTAAAGGCTCTGTCTGAATTTGACAGCCCTAATGATACA 3600  
 QY 1201 GluArgThrAsnIleGlnValThrValThrGlyProSerSerProSerPro----- 1217  
 Db 3601 GAAAGGACCAATATTCAGAGTGAACGCGTGAACGCGGCTTACCTCAACAGTCTGTAAAGTTT 3660

QY 1218 -----LeuAlaValAlaGlnPro 1223  
 Db 3661 CTGATTGACACACAAACCCCTTACTCTTCAAGACAGACAGAGCTTGTGGTATACGCCA 3720  
 QY 1224 MetAlaValAsnIleSerAlaAsnGlyPheGlyPheAlaIleCysGlnLeuAsnValVal 1243  
 Db 3721 ACCGCACTTAATTTTCCGCAAAATGTTTGGATTGCTAATTTTGACGCTCAATGTTGTA 3780  
 QY 1244 TyrAsnValLysAlaSerGlySerSerArgArgArgSerIleGlnAsnGlnGluAla 1263  
 Db 3781 TATATGTGAAGGCTTCTGGGCTTCTTCAAGAGCAAGATCTATCCAAATCAAGAAAGCC 3840  
 QY 1264 PheAspLeuAspValAlaValLysGluAsnLysAspPheLeuAsnHisValAspLeuAsn 1283  
 Db 3841 TTTGATTTAATGATGTTGCTGTAAAGAAATTAAGATCATCTCAATCATCTGTGGATTGAAAT 3900  
 QY 1284 ValCysThrSerPheSerGlyProGlyArgSerGlyMetAlaLeuMetGluValAsnLeu 1303  
 Db 3901 GTGTGTACAAAGCTTTTCCGGCCCGGGTAGAGTGCATGCTCTTATGAGAAATTAACCTTA 3960  
 QY 1304 LeuSerGlyPheMetValProSerGluAlaIleSerLeuSerGluThrValLysLysVal 1323  
 Db 3961 TTTAAGTGGCTTTATATGTCCTTCAAGAAACAAATTTCTGACGACAGACAGTGAAGAAAGTG 4020  
 QY 1324 GluTyrAspHisGlyLysLeuAsnLeuTyrLeuAspSerValAsnGluThrGlnPheCys 1343  
 Db 4021 GAATATGATCATGGAACAACTCAACCTTATTTAGATTTCTGTAAATGAACCCAGTTTGT 4080  
 QY 1344 ValAsnIleProAlaValAlaGlnAsnPheLysValSerAsnThrGlnAsnAlaSerValSer 1363  
 Db 4081 GTTAATATTTCTCTGTGAAGAAACCTTAAAGTTTCAAAATCCCAAGATGCTTCAAGTCTTC 4140  
 QY 1364 IleValAspTyrTyrGluProArgArgGlnAlaValArgSerTyrAsnSerGluValLys 1383  
 Db 4141 ATATGATTAATCTATATGAGCCAAAGACAGACAGCCGCTGAGAAATTCATCTGAAATGAAG 4200  
 QY 1384 LeuSerSerCysAspLeuCysSerAspValGlnGlyCysArgProCysGluAspGlyAla 1403  
 Db 4201 CTGTCTCTCTGTACCTTTGCAAGTATGTCACAGGCTGCGCTCTGTGAGATGAGACT 4260  
 QY 1404 SerGlySerHisHisHisSerSerValIlePheIlePheCysPheLysLeuLeuTyrPhe 1423  
 Db 4261 TCAGGCTCCCATATACATCTTCAAGTCATTTATTTCTGTTTCAAGCTTCTGTACTTT 4320  
 QY 1424 MetGluLeuTrpLeu 1428  
 Db 4321 ATGGAATTTGGCTG 4335  
 Db  
 RESULT 11  
 AAA62010  
 ID AAA62010 standard; DNA; 4473 BP.  
 AC AAA62010;  
 XX  
 DT 02-FEB-2001 (first entry)  
 XX  
 DE Hydrophobic domain protein cDNA HP02837 isolated from HT-1080 cells.  
 KW Human; secreted protein; membrane protein; hydrophobic domain;  
 KW proliferation control; differentiation induction; material transport;  
 KW biophylaxis; signal receptor; ion channel; transporter; immunostimulant;  
 KW immunosuppressant; haematopoiesis regulator; chemokine; chemokine;  
 KW hemostasis; thrombolytic; anti-inflammatory; tumour inhibition;  
 KW autoimmune disease; Alzheimer's disease; Parkinson's disease; cancer;  
 KW gene therapy; ss.  
 OS Homo sapiens.  
 XX  
 OS  
 PN MO200029448-A2.  
 XX  
 PD 25-MAY-2000.  
 XX  
 PF 17-NOV-1999; 99MO-JP06412.



XX 17-NOV-1998; 98JP-0326255.  
 PR 22-DEC-1998; 98JP-0364315.  
 PR 16-MAR-1999; 99JP-0069811.  
 PR 27-APR-1999; 99JP-0119299.  
 PR 19-MAY-1999; 99JP-0138169.  
 XX  
 PA (SAGA ) SAGAMI CHEM RES. CENT.  
 PA (PROT-) PROTEGENE INC.  
 XX  
 PI Kato S, Kimura T;  
 XX  
 DR WPI, 2000-387753/33.  
 DR P-PSDB; AAB12127.  
 XX  
 PT Proteins comprising hydrophobic regions, such as secretory and membrane  
 PT proteins, useful in research and diagnostics and having various  
 PT activities e.g. immunomodulatory, antiinflammatory, chemokinetic,  
 XX hemostatic, thrombolytic -  
 XX  
 PS Claim 4; Page 253-261; 410pp; English.

CC Secretory proteins play important roles in the proliferation control, the  
 CC differentiation induction, the material transport and the biophysics of  
 CC cells. Membrane proteins have important roles as signal receptors, ion  
 CC channels and transporters. The present sequence is the coding sequence  
 CC for a human protein which has at least one hydrophobic domain. The  
 CC protein encoded by the present sequence may be a secretory or a membrane  
 CC protein. The encoded protein may have cytokine and cell  
 CC proliferation/differentiation activity, immune stimulating or suppressing  
 CC activity, haemopoiesis activity, tissue growth activity,  
 CC activating/inhibiting activity, chemotactic/chemokinetic activity, haemostatic  
 CC and thrombolytic activity, anti-inflammatory activity and tumour  
 CC inhibition activity. The present sequence could therefore be used for  
 CC treatment of autoimmune disease, Alzheimer's disease, Parkinson's  
 CC disease, and cancer via gene therapy.

XX Sequence 4473 BP; 1328 A; 867 C; 952 G; 1326 T; 0 other;

#### Alignment Scores:

Pred. No.: 0 Length: 4473  
 Score: 7307.50 Matches: 1424  
 Percent Similarity: 98.62% Conservative: 1  
 Best Local Similarity: 98.55% Mismatches: 3  
 Query Match: 99.45% Indels: 17  
 DB: 21 Gaps: 1

US-10-020-095-4 (1-1428) x AAA62010 (1-4473)

QY 1 MetGlnGlyProProLeuLeuThrAlaAlaHisLeuLeuCyValCyThrAlaAlaLeu 20  
 DB 45 ATGCAAGGCGCCACCGCTCTGACCGCGCCACCTCTCTGCGTGGACCGCGCGCGTG 104  
 QY 21 AlaValAlaProGlyProArgPheLeuValThrAlaProGlyIleIleArgProGly 40  
 DB 105 GCGTGGCTCCCGGCGCTCGGTTTCTGTGACGAGCCCAAGGATCATCAGCGCCGAGGA 164  
 QY 41 AenValThrIleGlyValGluLeuLeuGluHisCysProSerGlnValThrValIysAla 60  
 DB 165 AATGTACTATTGGGGGTGAGGCTTCGGAACACTGCGCTTTCACAGGTGACTGTGAAGGGG 224  
 QY 61 GluLeuLeuThrAlaSerAenLeuThrValSerValLeuGluIleGluGlyValIlePhe 80  
 DB 225 GAGCTGCTCAAGACACATCAACCTCACTGCTCTGCTCTGGAACGAGAGTCTTT 284  
 QY 81 GluIleGlySerPheIleThrLeuThrLeuProSerLeuProLeuAenSerAlaAepGlu 100  
 DB 285 GAAAAAGGCTCTTTAAGACACTTACTCTTCATCACTACCTCTGAACAGTGCAGATGAG 344  
 QY 101 IleTyrGluLeuArgValThrGlyAgtThrGlnAepGluIleLeuPheSerAenSerThr 120  
 DB 345 ATTATGACTAGTGTAAACCGGACGTACCCAGATGAGATTATTTCTTAATAGTACC 404

QY 121 ArgLeuSerPheGluThrLysArgIleSerValPheIleGlnThrAspValaLeuTyr 140  
 DB 405 CGCTTATATCTTGGAGCCAAAGAAATATCTGTCTTCAATTCAAACAGACAGGCTTATAC 464  
 QY 141 LysProLysGlnGluValLysPheArgIleValThrLeuPheSerAspPheLysProTyr 160  
 DB 465 AAGCCAAAGCAAGAGTGAAGTTGCAATGTGTACACTCTCTCCAGATTTTAAGCCTTAC 524  
 QY 161 LysThrSerLeuAenIleLeuIleLysAepProLysSerAenLeuIleGlnGlnThrLeu 180  
 DB 525 AAAACCTCTTTAAACATTTCTCAATTAAGAGACCCCAATCAAAATTTAATCAACAGTGTG 584  
 QY 181 SerGlnGlnSerAspLeuGlyValIleSerLysThrPheGlnLeuSerSerHisProIle 200  
 DB 585 TCACAAACAAGATGATCTTGAGACTCTTTCACAAAACCTTTACGCTATCTTCCATCCATAC 644  
 QY 201 LeuGlyAepTyrSerIleGlnValGlnValAenAepGlnThrTyrTyrGlnSerPheGln 220  
 DB 645 CTTGGTGACTGGTCTATTCAGTTCAAGTGAATGACCAACATATTAATCAATCAATTCAG 704  
 QY 221 ValSerGluTyrValLeuProLysPheGluValThrLeuGlnThrProLeuTyrCysSer 240  
 DB 705 GTTTCAGAAATATGATATTAACCAAAATTTGAAGTACTTGCAGACACCATTAATATGTTCT 764  
 QY 241 MetAenSerLysHisLeuAenGlyThrIleThrAlaLysTyrThrTyrGlyLysProVal 260  
 DB 765 ATGATTTCTAAGCATTTTAATATGTCACATCAAGCAAGATATACATATGCAAGCAGCTG 824  
 QY 261 LysGlyAepValThrLeuThrPheLeuProLeuSerPheTyrGlyLysLysAenIle 280  
 DB 825 AAAGAGAGCTAAACCTTACATTTTACCTTATCTTTGGGGAAGAAAGAAATATTT 884  
 QY 281 ThrLysThrPheLysIleAsnGlySerAlaAenPheSerPheAenAepGluIleMetLys 300  
 DB 885 ACAAACAATTTAAATTAATGAATCTGCAACCTTCTTTAAATGAATCAAGATGAGAA 944  
 QY 301 AenValMetAepSerSerAsnGlyLeuSerGluTyrLeuAepLeuSerProGlyPro 320  
 DB 945 AATGTAAATGATTTCTTCAATGAGACTTCTGAAATCTCGATCTATCTTCCCTGAGACA 1004  
 QY 321 ValGluIleLeuThrThrValThrGluSerValThrGlyIleSerArgAenValSerThr 340  
 DB 1005 GTAGAAATTTTAAACAGAGTGCAGAAATCAGTTACAGGATTTCAAGAAATGTAAGCACT 1064  
 QY 341 AenValPhePheLysGlnHisAepTyrIleIleGluPhePheAepTyrThrThrValLeu 360  
 DB 1065 AATGTGTTCTCAACCAACATGATTAATCAATCAATTAATGATTAATCAATCAATCTCTG 1124  
 QY 361 LysProSerLeuAenPheThrAlaThrValLysValThrArgAlaAepGlyAenGlnLeu 380  
 DB 1125 AAGCAATCTCTCACTTCAACCTCAACGCACTGTGAAGGTAAGTCTGCTGATGAGCAACCACTG 1184  
 QY 381 ThrLeuGluGluArgArgAenAenValValIleThrValThrGlnArgAenTyrThrGlu 400  
 DB 1185 ACTCTTGAAGAAAGAAATTAATGATGATCAATCAAGTGCAGACAGAAATGATATCTGAG 1244  
 QY 401 TyrTyrSerGlySerAenSerGlyAenGlnLysPheGluIleGlnLysIleAenTyr 420  
 DB 1245 TACTGAGAGCGATCTTAACAGTGAATATCAAGAAATGAGAGCTGTTCAAGAAATTAATAT 1304  
 QY 421 ThrValProGlnSerGlyThrPheLysIleGluPheProIleLeuGluAepSerGlu 440  
 DB 1305 ACTGTCCTCCAAAGTGAACCTTTTAAGATTGAATTCCTCAATCTGAGAGATTCAGTGAAG 1364  
 QY 441 LeuGlnLeuLysAlaTyrPheLeuGlySerLysSerSerMetAlaValHisSerLeuPhe 460  
 DB 1365 CTACAGTTGAAGGCTTATTTCTTGTGTAAGAAAGTACATGAGCACTTATAGTGTGTT 1424  
 QY 461 LysSerProSerLysThrTyrIleGlnLeuLysThrArgAepGluAenIleLysValGly 480  
 DB 1425 AAGTCTCTAGTAAGACATATCACTCAAAACCAAGATGAGAAATTAATTAAGTGTGGA 1484  
 QY 481 SerProPheGluLeuValValSerClyAenLysAenLysArgLeuLysGluLeuSerTyrMetVal 500

Db	2555	GCTTCGATGCTATCA	CCCAAGATGATTTT	AGTAAAGCTGA	AGGAATGAAAAATCATAT	2622	
Qy	861	SerGlnSerIleLeu	LeuAspLeuThr	AspAsnArgLeuGlnSer	ThrLeuLeuThrLeu	880	
Db	2625	TCACATCATCATTTA	TAGACTTGACATG	CAATAGGCTACAGAGTACCTG	AAAAAATTGG	2688	
Qy	881	SerPheSerPhe	ProProAsnThr	ValThrGlySerGlu	ArgValGlnIleThrAlaIle	900	
Db	2685	AGTTCTCATTTCC	CTCTTAATACAGTGC	CTGGCAGTGAAAGATTC	CAATCATCTCAATT	2744	
Qy	901	GlyAspValLeu	GlyProSerIle	AsnGlyLeuAlaSerLeu	IleArgMetProTyrGly	920	
Db	2745	GGAGATGTTCTTGG	CTCTTCATCAATGGCT	TACGCTCATTTGATTTG	CGATTCCTTATGGC	2804	
Qy	921	CysGlyGln	AsnMetIleAsn	PheAlaProAsnIleTyr	IleLeuAspTyrLeuThr	940	
Db	2805	TGTGTGAAACA	GAGACATGATAAAT	TTTGTCTCCAATATTT	TTCATTTTGATTTATCTGACT	2864	
Qy	941	LeuLeuLeuGlnLeu	ThrAspAsnLeu	LeuGlyAlaLeuSerPhe	MetArgGlnGly	960	
Db	2865	AAAAAGAAACA	CTGACAGATTAATT	TGAAAAAAAGCTCTT	CAATTATAGCGCAAGCT	2922	
Qy	961	TyrGlnArgGlu	LeuLeuLeuTyrGln	ArgGlnAspGlySerPhe	SerAlaPheGlyAsnTyr	980	
Db	2925	TACCAAGAGAACT	CTCTTATCAGAGGGA	AAATGGCTTTTCACTG	CTCTTTGGGAAATAT	2988	
Qy	981	AspProSerGly	SerThrTyrLeuSer	AlaPheValLeuArgCys	PheLeuGluAlaAsp	1000	
Db	2985	GACCTTCTGGG	AGACACTTGGTGTG	CAGCTTTTGTTTAAAT	GTGTTCTCTTGAAGCCGAT	3044	
Qy	1001	ProTyrIle	AspIleAspGln	AsnValIleuHisArg	ThrTyrThrTyrLeu	LeuGlyHis	1020
Db	3045	CCTTACATAGAT	TATTTGATCAGAAAT	GTGTTCACACAGAACAT	CACCTGCTTAAAGACAT	3104	
Qy	1021	GlnLeuSerSer	AsnGlyGluPhe	ThrAspProGlyArg	ValIleHisSerGlnLeu	GlnGly	1044
Db	3105	CAGAAATCAAC	GGGAAATTTTGGAT	TCACAGAAAGAGATTC	ATGAGAGCTTCAAGCT	3164	
Qy	1041	GlyAsnLeu	SerProValThr	LeuThrAlaTyrIle	ValIleThrSerLeuLeu	GlyTyrArg	1066
Db	3165	GGCAATAAAA	ATCCAGTACACTT	TACAGCCATATTTG	TAATTTCTCTCTGGGATATG	3222	
Qy	1061	LysTyrGln	ProAsnIleAsp	ValGlnGluSerIle	HisPheLeuGluSerGlu	PheSer	1086
Db	3225	AAGTATCAG	CTTAACATGATG	TGTCAGAGCTCAT	CTCTTTTGGAGCTG	GAATTCAGT	3284
Qy	1081	ArgGlyIle	SerAspAsnTyr	ThrLeuAlaLeuIle	ThrTyrAlaLeuSerSer	ValGly	1100
Db	3285	AGAGGAATTT	CAGCAATTAATCT	TACCCCTTAATTA	TGCAATGTCACTAC	TGCTGAGTGGG	3344
Qy	1101	SerProLeu	SerAlaValAlaLeu	AsnMetLeuThrTyr	ArgAlaGlnGlnGly	Gly	1120
Db	3345	AGTCTTAAG	CGAAGAACTTTGA	ATATGCTGACTTGA	AGACAGAACAAAGAA	AGGTGGC	3404
Qy	1121	MetGlnPhe	TyrValSerSer	GluSerIleLeuSer	AspSerTyrGlnPro	ArgSerLeu	1140
Db	3405	ATGCATATCT	GGGTGTCATCAG	AGTCCAACTTTCTG	CACTCTGCGACGCC	CGCTCCTG	3464
Qy	1141	AspIleGlu	ValAlaAlaTyr	AlaLeuLeuSerHis	PheLeuGlnPheGln	ThrSerLeu	1166
Db	3465	GATATTTGA	GTGAGCTTATG	CACTGCTTCA	CACTTTCAATTC	CAACATCTTGAG	3522
Qy	1161	GlyIlePro	IleMetArgTyr	LeuSerArgGln	ArgAsnSerLeuGly	GlyPheAlaSer	1188
Db	3525	GGAAATCCAA	TTTAAGAGGTGCT	TAAAGCAAGGCA	AGAAATACCTTGG	TGTGGTTTGCACT	3588
Qy	1181	ThrGlnAsp	ThrThrValAlaLeu	AlaLeuSerGlu	PheAlaAlaLeuMet	AsnThr	1200
Db	3585	ACTCAGATTA	CACTGTGCTTTAA	AGGCTGTGTCTGA	ATTTTGACGCC	CTATATGATATCA	3644
Qy	1201	GluArgThr	AsnIleGln	ValIleThrVal	ThrGlyProSerSer	ProSerPro-----	1217
Db	3645	GAAGAGCA	CAAAATATCAAG	TGACCGGTG	CGGGCTTGTGCTAC	CAAGTCTCTGTAAAGTTT	3704

QY 1218 -----Leu1aVal1aGlnPro 1223  
 Db 3705 CTGATTGACACACACACCGCTTACTCTTCAGACAGCAGAGCTTCTGTGTACAGCCA 3764  
 QY 1224 MetAlaValAsn1LeSerAlaanglyPheGlyPheAla1LeCysGlnLeuVal1Val 1243  
 Db 3765 ACCGCACTTAATATTTCCCGCAAAVGGTTTGGATTGTCTATTTGTCACTCAATGTGTA 3824  
 QY 1244 TyrAsnVal1aValaSerGlySerSerArgArgArgSer1LeGlnGlnGlnGln1a 1263  
 Db 3825 TATTAATGTAAGGCTTCTGGGCTTCTAGAGACGAAGATCTATTCAAATCAAGAGCC 3884  
 QY 1264 PheAspLeuAspVal1aVal1aGlyAsn1aAspAspLeuAsn1aVal1aAspLeuAsn 1283  
 Db 3885 TTTGATTATGATCTCTCTGTAAAGAAATTAAGATCTCAATCATGTGATTGAAAT 3944  
 QY 1284 Val1CysThrSerPheSerGlyProGlyArgSerGlyMetAlaLeuMetGluValaLeu 1303  
 Db 3945 GTGTGTACAAGCTTTTCGGGCGCGGTAGAGATGCGCTCTTATGGAAGTTAACCTA 4004  
 QY 1304 LeuSerGlyPheMetValProSerGluAla1LeSerLeuSerGluThrVal1aLysVal 1323  
 Db 4005 TTTAAGTGGCTTATGATGCTTCAGAGCAATTTCTCTGACGACAGCACTGAAGAAAGTG 4064  
 QY 1324 GluTyrAspHisGlyLysLeuAsnLeuTyrLeuAspSerValaAsnGlnThrGlnPheCys 1343  
 Db 4065 GAATATGATCATGGAAGAACTCAACCTCTATTAGATTCTGTAAATTAACCCAGTTTGT 4124  
 QY 1344 ValaAsn1LeProAlaValaArgaAspPheVal1SerAsnThrGlnaAsp1aSerValaSer 1363  
 Db 4125 GTTAATATTCCTGCTGTGAAGAACTTAAAGTTTCAAAATACCAAGATGCTTCAGTGTC 4184  
 QY 1364 IleValAspTyrTyrGlnProArgaGlnAlaValaArgSerTyrAsnSerGluVal1aLys 1383  
 Db 4185 ATTAGTGATTACTATATAGCCAAAGACAGCGCGTGAAGATTTACAACTCTGAAGTGAAG 4244  
 QY 1384 LeuSerSerCysAspLeuCysSerAspVala1GlnGlyCysArgProCysGluAspGlyAla 1403  
 Db 4245 CTGTCTCTCTGTACCTTTTGCAATGATGTCACAGGCTGCGCTGTGTGAGAGTGAAGCT 4304  
 QY 1404 SerGlySerHisHisSerSerVal1aPhe1aPheCysPheLysLeuLeuTyrPhe 1423  
 Db 4305 TTAGGCTCCCATCATCACTTTCAGTCATTTTATTTCTGTTCAGGCTTCGTACTTT 4364  
 QY 1424 MetGluLeuTyrPhe 1428  
 Db 4365 ATGGAATTTGGCTG 4379  
 RESULT 12  
 AAD49436  
 ID AAD49436 standard; DNA; 4146 BP.  
 AC AAD49436;  
 XX 24-MAR-2003 (first entry)  
 DE Human blood cell surface antigen, CD109 encoding DNA #1.  
 XX Human; GPI-anchored TGF-beta1 binding protein; r150 protein; cancer;  
 KW glycosylphosphatidylinositol; transforming growth factor-beta1;  
 KM therapy; blood cell surface antigen; CD109; gene; da.  
 XX Homo sapiens.  
 OS  
 XX  
 XX Key Location/Qualifiers  
 FH 1..4146  
 FT CDS.  
 FT /tag= a  
 FT /product= "Human CD109 protein"  
 FT /transl\_except= (pos:2044..2046, aa:Xaa)  
 FT /note= "Xaa corresponds to Ser, Tyr; No start  
 FT and stop codon"  
 FT /partial

XX PN WO200265942-A2.  
 XX PD 31-OCT-2002.  
 XX PF 24-APR-2002; 2002WO-CA00560.  
 XX PR 24-APR-2001; 2001US-285713P.  
 XX PR 14-FEB-2002; 2002US-356163P.  
 XX PA (UWMC-) UNIV MCGILL.  
 XX PI Philip A, Tam B;  
 XX DR WPI: 2003-093100/08.  
 XX DR P-PDB; MAE32014.  
 XX PT Novel transforming growth factor (TGF)-beta 1 binding reagent which  
 PT comprises r150 protein which acts as accessory receptor of TGF-beta,  
 PT useful for negatively modulating TGF-beta activity, and thus for  
 PT treating cancer  
 XX PS Claim 7, Page 109-112; 127pp; English.  
 CC The invention relates to novel transforming growth factor (TGF)-beta1  
 CC binding reagent which comprises a glycosylphosphatidylinositol (GPI)-  
 CC anchored TGF-beta1 binding protein referred to as r150 which acts as  
 CC an accessory receptor of TGF-beta. The invention is used for negatively  
 CC modulating TGF-beta activity, and thus for treating conditions  
 CC characterized by overproduction of TGF-beta, such as cancer. Antisense  
 CC molecules of the invention are used for increasing TGF-beta availability  
 CC and increase graft success. The present sequence is human blood cell  
 CC surface antigen, CD109 encoding DNA. CD109 is a variant of r150 sequence.  
 XX SQ Sequence 4146 BP; 1260 A; 776 C; 885 G; 1224 T; 1 other;  
 Alignment Scores:  
 Pred. No.: 0 Length: 4146  
 Score: 7081.00 Matches: 1379  
 Percent Similarity: 99.86% Conservative: 1  
 Best Local Similarity: 99.78% Mismatches: 2  
 Query Match: 96.37% Indels: 0  
 DB: Gaps: 0  
 US-10-020-095-4 (1-1428) x AAD49436 (1-4146)  
 QY 22 Vala1aProGlyProArgPheLeuVal1aThra1aProGly1Le1aArgProGly1aValaAsn 41  
 Db 1 GTGGCTCCCGGCGCTCGGTTTCTGTGTGACAGCCGAGGATCATCAGGCCCGAGGAAT 60  
 QY 42 ValThr1LeGlyVal1GlnLeuLeuGlnHisCysProSerGlnVal1ThrVal1aGlu 61  
 Db 61 GTGACTAATGGGGTGAAGCTTCTGTGAACATGCTCCCTTCAAGTGAATGTAAGCGGAG 120  
 QY 62 LeuLeuLysThrAlaSerAsnLeuThrValaSerVal1aGlnGlnGlnGln1aValaPheGlu 81  
 Db 121 CTGCTCAAGACGATCAACCTCACTCTCTGTGCTCGAAGACGAAGAGTCTTTGAA 180  
 QY 82 LysGlySerPheLysThrLeuThrLeuProSerLeuProLeuAsnSerAlaAspGluLe 101  
 Db 181 AAAGCTCTTTTAAGACCTTACTCTTCACTCACTCACTCACTCACTCACTCACTCACTCACT 240  
 QY 102 TyrGluLeuArgVal1aThrGlyArgThrGlnaAspGlu1LeuPheSerAsnSerThrArg 121  
 Db 241 TATGAGCTAAGTGTAAACCGGACGTACCCAGATGATTTATTTCTTAATATGATCCCGC 300  
 QY 122 LeuSerPheGluThrLysArg1LeSerValaPhe1LeGlnThrAspLysAlaLeuTyrLys 141  
 Db 301 TTATATTTTGAACCAAGAAATATCTGTCTTCACTTAACCAAGACAGAGCCCTTATACAG 360  
 QY 142 ProLysGlnGlnVal1aLysPheArg1LeVal1aThrLeuPheSerAspPheLysProTyrLys 161  
 Db 361 CCAAGCAAGAAAGTGAAGTTTTCGATTTGTTCACCTCTTCACTCTTCAAGATTTTAAGCTTACAA 420

QY 162 ThrsleuAenlleuileuysapProlyserAenlleuileglnlntprleuSer 181  
DB ACCTTTAAACATTCATTAGAGACCCTAAATTTGATCCAAACAGTGGTTC 480  
QY 182 GlnGlnSerApleuGlyValleuSerlyThrPheGlnleuSerSerHieProleu 201  
DB CAACAAAGTGAATCTTGAGTCAATTCCTTCCATCTTCCATCCAAATCTT 540  
QY 202 GlyAspTrpSerileGlnValGlnValAsnAspGlnThrTyrrGlnSerPheGlnVal 221  
DB GATGCTGCTCATTCAGTTCAGAGTGAATGACAGACATCTCATCAATTCAGGTT 600  
QY 222 SerGlnTyrrValleuProlyserPheGlnValThrleuGlnThrProleuTyrrCysSerMet 241  
DB TCAGATATGATATACCAAAATTTGAGTGAATTCGACAGACACATTTATTTGTTATG 660  
QY 242 AsnSerlyHieAsnleuGlyThrleThrAlaTyrrTyrrGlyLysProValLys 261  
DB AATCTTAAGCATTTTAAATGATACCATCAAGGCAAAAGTATGAGTGAACCCAGTGA 720  
QY 262 GlyAspValThrleuThrPheleuProleuSerPheTrpGlyLysLysAsnilleThr 281  
DB GAGAGCTGACGCTTACATTTTACCTTTATCTTTGGGAAAGAGAAATATTTACA 780  
QY 282 LysThrPheLysIleAsnGlySerAlaAsnPheSerPheAsnAspGlnGlnMetLysAsn 301  
DB AAAACATTTAAGATTAATGATCTGCAAACTTCTTTATGATGAAGACATGAAAAAT 840  
QY 302 ValMetAspSerSerAsnGlyLeuSerGlnTyrrleuAspLeuSerSerProGlyProVal 321  
DB GTATATGATCTTCAATGAGCTTTCGAAATACCTGATCTATCTTCCCTGGACCACTA 900  
QY 322 GlnleuLeuThrThrValThrGlnSerValThrGlyIleSerArgAsnValSerThrAsn 341  
DB GAAATTTTAACACAGTACAGAAATCACTTACAGATTTTCAAGAAATGAAGACATAT 960  
QY 342 ValPhePheLysGlnHieAspTyrrIleleGlnPhePheAspTyrrThrThrValleuLys 361  
DB GTGTTCTTCAAGCAACATGATTAATCATTTGATTTTGTATATACATCTGCTTGAAG 1020  
QY 362 ProSerleuAsnPheThrAlaThrValLysValThrArgAlaAspGlyAsnGlnleuThr 381  
DB CCATCTCTCAACTTCAACGACCTGTGAAGGTAACTCTGCTGATGACCAACTGACT 1080  
QY 382 LeuGlnGlnArgAsnAsnValIleThrValThrGlnArgAsnTyrrThrGlnTyrr 401  
DB CTTGAAAGAAAGAAATGATGATGATCAATCAAGTACACAGACAGAAATCATATGAGTAC 1140  
QY 402 TrpSerGlySerAsnSerGlyAsnGlnLysMetGlnAlaValGlnLysIleAsnTyrrThr 421  
DB TGGAGCGGATCTTAACAGTGAATTCAGAAATGGAAGCTGTTCCAGAAATTAATTAATCT 1200  
QY 422 ValProGlnSerGlyThrPheLysIleGlnPheProIleleuGlnAspSerSerGlnleu 441  
DB GTCCCCCAAGTGAATCTTTAAGATTAATTTCCCAATCTCGAGAGATTTCCAGAGAGCTA 1260  
QY 442 GlnleuLysAlaTyrrPheleuGlySerlySerSerMetAlaValHieSerleuPheLys 461  
DB CAGTTGAAGGCTTATCTTCTGTAGTAAAGTGAAGTGAAGTCAATGATCTTTTAAG 1320  
QY 462 SerProSerlyThrTyrrIleGlnleuLysThrArgAspGlnAsnIleLysValGlySer 481  
DB TCTCTTAAGTGAATCAATCACTCAACTAAACCAAGATGAAATTAATTAAGTGGAGTGC 1380  
QY 482 ProPheGlnleuValIleSerGlyAsnLysArgleuLysGlnleuSerTyrrMetValVal 501  
DB CCTTTTGAAGTGGTGTAGTGGCAACAAAGATTTGAAGAGTAAAGTATATATGATAGTA 1440  
QY 502 SerArgGlyGlnleuValAlaValGlyLysGlnAsnSerThrMetPheSerleuThrPro 521  
DB TCCAGGGGACAGTGTGGCTGTGTGAGAAACAAATTCACAAATGTTCTCTTTAAACACA 1500

QY 522 GlnAsnSerTrpThrProlysaAlaCysValIleValTyrrIleGlnAspAspGlyGln 541  
DB GAAATTTCTGACCTCCAAAAGCTGTGATTTGTGATATTAATGTAAGATGAGGGAA 1560  
QY 542 IleIleSerAspValleuLysIleProValGlnleuValPheLysAsnLysIleLysleu 561  
DB ATTATAGATGATGTCTTAAATATCTGTCACCTGTGTTTAAATAGATGAAGCTTA 1620  
QY 562 TyrrTrpSerlyValLysAlaGlnProSerGlnLysValSerleuArgIleSerValThr 581  
DB TATTTAGATTAAGTGAAGCTGAAACCATCTGAGAAAGTCTCTTTAGATCTCTGACACA 1680  
QY 582 GlnProAspSerIleValGlyIleValAlaValAspLysSerValAsnleuMetAsnAla 601  
DB CAGCTTACTCCATGATGTTGGATTTAGCTGTGACAAAGTGAATCGAATCGAATGCC 1740  
QY 602 SerAsnAspIleThrMetGlnAsnValIleValGlnleuGlnleuTyrrAsnThrGlyTyrr 621  
DB TCTATATATATTAACAATGAAATGTGTCCATGAGTGAACCTTATTAACACAGATAT 1800  
QY 622 TyrrleuGlyMetPheMetAsnSerPheAlaValPheGlnGlnLysCysGlyleuThrValleu 641  
DB TATTTAGCATCTTTCATGAATCTTTTGCAGTCTTTCAAGAAATGTGACTCTGGTATTTG 1860  
QY 642 ThrAspAlaAsnleuThrlyAspTyrrIleAspGlyValTyrrAspAsnAlaGlnTyrrAla 661  
DB ACAGATCAAACTCCAGAAAGATTAATGATGAGTGTATTAACAATGACAGAAATATGCT 1920  
QY 662 GlnArgPheMetGlnGlnAsnGlnGlyHieIleValAspIleHieAspPheSerleuGly 681  
DB GAGAGGTTTATGAGAGAAATGAAAGACATATGATGATTAATGATCACTTTCTTTGGCT 1980  
QY 682 SerSerProHieValArgLysHiePheProGlnThrTrpIleThrleuAspThrAsnMet 701  
DB AGCAGTCCACATGCTCCGAAAGCATTTTCCAGACATTTGATGCTGACAGACCAACATG 2040  
QY 702 GlyTyrrArgIleTyrrGlnleuPheGlnValThrValProAspSerIleThrSerTrpVal 721  
DB GGTTCAGATTTATACCAAGATTTGAAGTATGTAAGTCACTGATCTTCACTCTTGGGCTG 2100  
QY 722 AlaThrGlyPheValIleSerGlyAspLeuGlyLysleuThrThrThrProValGln 741  
DB GCTATCGGTTTGTGATCTCTGAGACCTGGGCTTGGACTTAACAATCACTCAAGAGAG 2160  
QY 742 LeuGlnAlaPheGlnProPhePheIlePheLeuAsnleuProTyrrSerValIleArgGly 761  
DB CTCAGACCTTCCACACATTTTTCATTTTGAATCTTCCACTCTGTTACAGAGGT 2220  
QY 762 GlnGlnPheAlaLeuGlnIleThrIlePheAsnTyrrleuLysAspAlaThrGlnValLys 781  
DB GAAAGATTTGCTTGGAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAG 2280  
QY 782 ValIleIleGlnLysSerAspLysPheAspIleleuMetThrSerSerGlnIleAsnAla 801  
DB GATATCATTTGAGAAAGTGAACAAATTTGATATGATGATCAATCAATGAATGAATATGCC 2240  
QY 802 ThrGlyHieGlnGlnThrleuLysValProSerGlnAspGlyAlaThrValleuPhePro 821  
DB ACHAGCCACAGACAGACCTTCTGTGTTCCAGTGAAGATGGGCACTGCTCTTTTCCC 2400  
QY 822 IleArgProThrHieAsnGlyGlnIleProIleThrValThrAlaLeuSerProThrAla 841  
DB ATCAGGCAACACATCTGGGAAATTTCCATCAAGTCAACACTCTTCAACCACTGCT 2460  
QY 842 SerAspAlaValThrGlnMetIleleuValLysAlaGlnGlyLysIleGlnLysSerTyrrSer 861  
DB TCTGATGCTGTCAACCAAGATTAATTAAGGCTGAAGGAATGAAGAAATCATATTTCA 2520  
QY 862 GlnSerIleleuLysAspLeuThrAspAsnArgleuGlnSerThrleuLysThrleuSer 881  
DB CAATCAATCTTATTAAGCTTGAACAAATGAGCTCAAGATACCTTGAAATCTTGAAGT 2580  
QY 882 PheSerPheProAsnThrValThrGlySerGlnValGlnIleThrAlaIleGly 901

```

Db      2581  TTCTCATTTCTCTCATATACAGTACGTCGACGAGAAAGATTCAATTCATCGAATTTGAA 2640
Qy      902   AspValLeuGlyProSerIleAsnGlyLeuAlaSerLeuIleArgMetProTyrGlyCys 921
Db      2641  GATGTTCTTGTCCTTCCATTCATAGGCTTACGCTCATTTGATTCGATGCTTATGCTCT 2700
Qy      922   GlyGluGlnAsnMetIleAsnPheAlaProAsnIleTyrIleLeuAspTyrLeuThrIys 941
Db      2701  GGTGAACAGACATGATTAATTTTGCTCAAAATATTACATTTTGATTTATCGACTAAA 2760
Qy      942   LysIleGlnLeuThrAspAsnLeuGlyIleValAlaLeuSerPheMetArgGlnGlyTyr 961
Db      2761  AAGAAACAACCTGACACATTAATTTGAAAAGAAAGCTTTTCATTTATGAGGCAAGCTTAC 2820
Qy      962   GlnArgGluLeuLeuTyrGlnArgGluAspGlySerPheSerAlaPheGlyAsnTyrAsp 981
Db      2821  CAGAGAAACCTTCTCATACAGAGGAAAGTGGCTTTTCAGTGCCTTTTGGAATTAATGAC 2880
Qy      982   ProSerGlySerThrTyrLeuSerAlaPheValLeuArgCysPheLeuGluAlaAspPro 1001
Db      2881  CCTTCTGGGAGCCTTGTTGTCAGCTTTGTTTAAAGATGTTTCTTGAAGCCGATCCT 2940
Qy      1002  TyrIleAspIleAspGlnAsnValLeuHisArgGlnTyrThrTyrLeuIysGlyHisGln 1021
Db      2941  TACATGATATTGATCAGATGATGTTACACAGAACATACACTTGCTTAAAGCACTCAG 3000
Qy      1022  LysSerAsnGlyGluPheTyrAspProGlyArgValIleHisSerGluLeuGlnGlyIys 1041
Db      3001  AATATCAACGGTGAATTTTGGATCCAGAAAGGTATTCATATGAGGTTCAAGGTGGCC 3060
Qy      1042  AsnLysSerProValThrLeuThrAlaTyrIleValThrSerLeuLeuGlyTyrArgIys 1061
Db      3061  AATTAAGTCCAGTAACTTACCTTACAGCCTATATTGTACTTCTCTGGGATATGAAAG 3120
Qy      1062  TyrGlnProAsnIleAspValGlnGlySerIleHisPheLeuGlnSerGluPheSerArg 1081
Db      3121  TATCAGCCTTAACATTGATGTCAGAGTCTATCCATTTTGGAGTCAATTCAGTACATA 3180
Qy      1082  GlyIleSerAspAsnTyrThrLeuAlaLeuIleThrTyrAlaLeuSerSerValGlySer 1101
Db      3181  GGAATTTCAAGACATTAATCTTACGCCCTTATTAACCTTATGCAATGTCATCAGTGGGAG 3240
Qy      1102  ProLysAlaLysGluAlaLeuAsnMetLeuThrTyrArgAlaGluGlnGlyIysMet 1121
Db      3241  CCTAAAGCAGAGAACTTTGAATATGCTGACTTGAGAGCAGACAAAGAGTGGCATTG 3300
Qy      1122  GlnPheTyrValSerSerGlnSerIysLeuSerAspSerTyrGlnProArgSerLeuAsp 1141
Db      3301  CAATTCGGGTGTCATCAGAGTCCAAACTTTTGACTCCCTGGCAGCCAGCTCCCTGGAT 3360
Qy      1142  IleGluValAlaAlaTyrAlaLeuLeuSerHisPheLeuGlnPheGlnThrSerGluGly 1161
Db      3361  ATTTGAAGTTGACGCCCTATGCTACCTGCTCACTTCTTCAATTTTCAGACTTCTGAGGGA 3420
Qy      1162  IleProIleMetArgTyrLeuSerArgGlnArgAsnSerLeuGlyGlyPheAlaSerThr 1181
Db      3421  ATCCCAATTATGAGTGGCTTACAGAGGCAAAAGAAATGCTGGGTGCTTTTGCAATTAAT 3480
Qy      1182  GlnAspThrThrValAlaLeuLysAlaLeuSerGluPheAlaAlaLeuMetAsnThrGln 1201
Db      3481  CAGGATACCACTGTGCTTTAAAGCTCTGTCTGATTTTGCAAGCCCTAATGATACAGAA 3540
Qy      1202  ArgThrAsnIleGlnValThrValThrGlyProSerSerProSerProLeuAlaValVal 1221
Db      3541  AAGACAAATATCCAACTGACCGTGAAGGGGCTTACGCAAGTCTCTTGGCTGGGTGA 3600
Qy      1222  GlnProMetAlaValAsnIleSerAlaAsnGlyPheGlyPheAlaIleCysGlnLeuAsn 1241
Db      3601  CAGCCAAAGCAATTAATTTCCGCAAAATGTTTGGATTTGCTAATTTGTCAGCTCAAT 3660
Qy      1242  ValValTyrAsnValLysAlaSerGlySerSerArgArgArgSerIleGlnAsnGln 1261

```

```

Db      3661  GTTGTATATATATGAGGCTTGGGTCTTCTTACAGACAGAAATCTATCCAAATCA 3720
Qy      1262  GluAlaPheAspLeuAspValAlaValIysGluAsnLysAspAspLeuAsnHisValAsp 1281
Db      3721  GAGGCTTTGATTTTATGATGTTGCTGTATAAAGAAATTAAGATGATCTCAATCATGTGAT 3780
Qy      1282  LeuAsnValCysThrSerPheSerGlyProGlyArgSerGlyMetAlaLeuMetGluVal 1301
Db      3781  TTGAATGTGTACAAAGCTTTTCCGGCCCGGGTAGAGATGGCATGGCTTATGGAAGTT 3840
Qy      1302  AsnLeuLeuSerGlyPheMetValProSerGluAlaIleSerLeuSerGluThrValIys 1321
Db      3841  AACCTATTAAAGTGGTTTATGCTGCTTCAGAAAGCAATTTCTCGAGCGAGACAGTGAAG 3900
Qy      1322  LysValGluTyrAspHisGlyLysLeuAsnLeuTyrLeuAspSerValAsnGluThrGln 1341
Db      3901  AAGTGGAAATATGATCATGAGAAACTCAACTCTATTTAGATTTCTGTAAGAAACCAAG 3960
Qy      1342  PheCysValAsnIleProAlaValArgAsnPheLysValSerAsnThrGlnAspAlaSer 1361
Db      3961  TTTGTGTTAATATTCTGCTGTGAGAACTTTAAAGTTTCAAAATACCAAGATGCTTCA 4020
Qy      1362  ValSerIleValAspTyrTyrGluProArgArgGlnAlaValAlaGserTyrAsnSerGlu 1381
Db      4021  GTTCCATAGTGGATTACTATGAGCCAGAGACAGGCGGTGAGAAATTACCAACTCTGAA 4080
Qy      1382  ValLysLeuSerSerCysAspLeuCysSerAspValGlnGlyCysArgProCysGluAsp 1401
Db      4081  GTGAAGCTGTCTCTGAGACCTTTGACGATATGTCAGATATGCCAGGCTGCCGCTTGAGGAT 4140
Qy      1402  GlyAla 1403
Db      4141  GGAAGCT 4146

RESULT 13
AAD49437
ID AAD49437 standard; DNA; 4197 BP.
XX
AC AAD49437;
XX
DT 24-MAR-2003 (first entry)
XX
DE Human r150 DNA #2.
XX
KW Human; GPI-anchored TGF-beta1 binding protein; r150 protein; cancer;
KW glycosylphosphatidyl inositol; transforming growth factor-beta1;
KW therapy; gene; db.
XX
OS Homo sapiens.
XX
FH Key 1.4197
FT CDS /tag= a
FT /product= "Human r150 protein #3"
FT /transl_except= (pos:2044..2046, aa:Xaa)
FT /note= "Xaa corresponds to Ser, Tyr; No start
FT and stop codon"
FT /partial
XX
XX WO200285942-A2.
XX
XX 31-OCT-2002.
XX
XX 24-APR-2002; 2002WO-CA00560.
XX
XX 24-APR-2001; 2001US-285713P.
XX
XX 14-FEB-2002; 2002US-356163P.
XX
XX (UWMC-) UNIV MCGILL.
XX
XX Phillip A, Tam B;
XX
XX WPI; 2003-093100/08.
DR

```

DR P-PSDB; AAB32015.

XX Novel transforming growth factor (TGF)-beta 1 binding reagent which  
PT comprises r150 protein which acts as accessory receptor of TGF-beta,  
PT useful for negatively modulating TGF-beta activity, and thus for  
PT treating cancer -

XX Claim 7; Page 118-120; 127pp; English.

XX The invention relates to novel transforming growth factor (TGF)-beta1  
CC binding reagent which comprises a glycosylphosphatidylinositol (GPI)-  
CC anchored TGF-beta1 binding protein referred to as r150 which acts as  
CC an accessory receptor of TGF-beta. The invention is used for negatively  
CC modulating TGF-beta activity, and thus for treating conditions  
CC characterised by overproduction of TGF-beta, such as cancer. Antisense  
CC molecules of the invention are used for increasing TGF-beta availability  
CC and increase graft success. The present sequence is human r150 protein  
CC encoding DNA.

XX Sequence 4197 BP; 1276 A; 788 C; 894 G; 1238 T; 1 other;

## Alignment Scores:

Pred. No.: 0 Length: 4197  
Score: 7068.50 Matches: 1380  
Percent Similarity: 98.71% Conservativity: 1  
Best Local Similarity: 98.64% Mismatches: 1  
Query Match: 96.20% Indels: 17  
DB: 25 Gaps: 1

US-10-020-095-4 (1-1428) x AAD49437 (1-4197)

QY 22 ValAlaProGlyIProArgPheLeuValThrAlaProGlyIleIleArgProGlyIValAsn 41  
Db 1 GTGGCTCCCGGGCCCTCGGTTTCTGGTGACAGCCCGAGGATCATCAGGCCCGAGGAAT 60  
QY 42 ValThrIleGlyValGluLeuLeuGluHisCysProSerGlnValThrValIleValGlu 61  
Db 61 GTGCTATTGGGGGTGAGACCTTCTGGACACAGCCCTTCACAGGAGTGTAAAGCGGAG 120  
QY 62 LeuLeuValThrAlaSerAsnLeuThrValSerValIleGluValIleGluValPheGlu 81  
Db 121 CTGGCTCAAGACAGCAATCACTCATCTCTGCTGCTGGAAGCAAGAGGCTCTTGAA 180  
QY 82 LysGlySerPheLysThrLeuThrLeuProSerLeuProLeuAsnSerAlaAspGluIle 101  
Db 181 AAAAGCTCTTTAAGACACTTACTCTTCATCATCACTCTGAAACAGGAGATGAGATT 240  
QY 102 TyrGluLeuArgValThrGlyValArgThrGlnAspGluIleLeuPheSerAsnSerThrArg 121  
Db 241 TATGAGCTACGTGTAAACCGAGACGTACCCAGATGAGATTATTCTTAATAGTACCAG 300  
QY 122 LeuSerPheGluThrValArgIleSerValPheIleGlnThrAspValIleLeuTyrLys 141  
Db 301 TTATCATTTTGAACCAAGAAATATCTGTCTTCAATCAACAGCAAGGCTTATACAG 360  
QY 142 ProLysGlnGluValIlePheArgIleValThrLeuPheSerAspPheLysProTyrLys 161  
Db 361 CCAAGCAAGAGTGAAGTTTCGCACTTTTACACTCTTCTCAAGATTTTAAGCTTACAA 420  
QY 162 ThrSerLeuAsnIleLeuIleLysAspProLysSerAsnLeuIleGlnGlnTyrLeuSer 181  
Db 421 ACCCTTTAAACATTTCTATTAGAGACCCCAATCAATTTGATCCAAACAGTGTGCA 480  
QY 182 GlnGlnSerAspLeuGluValIleSerIleThrPheGlnLeuSerSerHisProIleLeu 201  
Db 481 CAACAAGATGATCTTGAGTCAATTCCTTCCAACTTTTCCATCTTCCATCAATACTT 540  
QY 202 GlnAspTrpSerIleGlnValIleValAsnAspGlnThrTyrArgIleSerPheGlnVal 221  
Db 541 GGTACCTGGTCTATTCAAGTTCAAGTGAATGACAGACATATTTCAATCACTTCCAGTT 600  
QY 222 SerGlnTyrValLeuProLysPheGluValThrLeuGlnThrProLeuTyrCysSerMet 241  
|||

Db 601 TCAGAAATATGATTAATCCAAAAATTTGAAGTACTTTGCAGACACCATTAATGTTCTATG 660  
QY 242 AsnSerLysHisLeuAsnGlyThrIleThrAlaLysTyrThrTyrGlyLysProValLys 261  
Db 661 AATTCTAGACATTTAAATGTTTACCATCAGCGAAAGTATCATATGGAAAGCCAGTGA 720  
QY 262 GlyAspValThrLeuThrPheLeuProLeuSerPheTrpGlyLysIleValAsnIleThr 281  
Db 721 GGAGACGTAAAGCTTCAATTTTACATTTTACCTTTATCTTTTGGGAAAGAAATATATTA 780  
QY 282 LysThrPheLysIleLeuGlnSerAlaAsnPheSerPheAsnAspGluGluMetLysAsn 301  
Db 781 AAAACATTTAAGATTAATGATGATTCGAAACTCTTTTATATATGAAGAGTAAAT 840  
QY 302 ValMetAspSerSerAsnGlyLysSerGlyTyrLeuAspLeuSerSerProGlyProVal 321  
Db 841 GTAAATGATTTCTCAATGATGATCTTCTGAATACCTGATATATCTTCCCTGAGCAGTA 900  
QY 322 GluIleLeuThrThrValThrGlnSerValThrGlyIleSerArgAsnValSerThrAsn 341  
Db 901 GAAATTTTAAACACAGTGAAGATCAGTTACAGTATTTCAAGAAATGTAAACACTAAT 960  
QY 342 ValPhePheLysGlnHisAspTyrIleIleGluPhePheAspTyrThrThrValIleLys 361  
Db 961 GTGTTCTTCAAGACACATGATTAATCATCATTTGATTTTGTATTATCTACTGTCTTGA 1020  
QY 362 ProSerLeuAsnPheThrAlaThrValIleValThrArgAlaAspGlyAsnGlnLeuThr 381  
Db 1021 CCACTTCTCAACTTTCACAGCCACTGTGAAGTAACTCGTCTGATGAGCAACCACTGACT 1080  
QY 382 LeuGluGluArgArgAsnValValIleThrValThrGlnArgAsnTyrThrGlyTyr 401  
Db 1081 CTGAAGAAAGAAATATATGATGATCATTAACAGTACACAGAAACTATACAGTAC 1140  
QY 402 TrpSerGlySerAsnSerGlyAsnGlnLysMetGluValIleGlnLysIleAsnTyrThr 421  
Db 1141 TGAAGCGATCTTAACAGTGAATCAAGAAATGAAAGCTGTTGAAGAAATTAATTA 1200  
QY 422 ValProGlnSerGlyThrPheLysIleGluPheProIleLeuGluAspSerSerGluLeu 441  
Db 1201 GTCCCCCAAGTGAAGCTTTTAAGATTTGAATTCCTCAATCTTGAGATTCAGATGACTA 1260  
QY 442 GlnLeuLysAlaTyrPheLeuGlySerLysSerSerMetAlaValHisLeuPheLys 461  
Db 1261 CAGTTAAGGCTTATTTCTTGGTAGTAAAGTGAAGATGAGCACTGATGCTGTTAAG 1320  
QY 462 SerProSerLysThrTyrIleGlnLeuLysThrArgAspGluAsnIleLysValIleGlySer 481  
Db 1321 TCTCTTAGTAAACATACATCACTCAATCAAAAGAGTGAAGAAATTAAGGTGGATCG 1380  
QY 482 ProPheGluLeuValIleSerGlyAsnLysArgLeuLysGluLeuSerTyrMetValIle 501  
Db 1381 CTTTGAAGTGGTGGTGTAGTGCACAAACGATTAAGATTAAGATTAAGTATGATGAT 1440  
QY 502 SerArgGlyGlnLeuValAlaValGlyLysGlnAsnSerThrMetPheSerLeuThrPro 521  
Db 1441 TCCAGGGGACAGTGGGCTGTAGGAAACAAATATCAACATGTTCTCTTAAACACCA 1500  
QY 522 GluAsnSerTrpThrProLysAlaCysValIleValTyrTyrIleGluAspAspGlyGlu 541  
Db 1501 GAAATTTCTTGGACTCCAAAGCCGTGTATGTTATTAATTAAGTATGAGTGGGAA 1560  
QY 542 IleIleSerAspValLeuLysIleProValGlnLeuValPheLysAsnLysIleLysLeu 561  
Db 1561 ATTATTAAGTATGATTTCTAAATAATCTCTGTTCACTGTTTAAATAATTAAGTAACT 1620  
QY 562 TyrTrpSerLysValIleLysAlaGluProSerGluLysValSerLeuArgIleSerValThr 581  
Db 1621 TATTGGAGTAAATGAAAGTGAACATCAGTGAAGAGTCTCTTGGATCTCTGTGACA 1680  
QY 582 GlnProAspSerIleValGlyIleValAlaValAspLysSerValAsnLeuMetAsnAla 601  
Db 1681 CAGCTGACTCCATATGTTGGATTTGATGCTGTGAACAAAGTGTGATCTGATGAATGCC 1740



QY	602	SeRsnApP1eThmEgLuAsnValVhIsgIuEugIuEuyrAaThhGlyTyr	621
Db	1741	TCtATGATATTACAATGGAATAAGTGTCCATGAGTGGACCTTAAACACAGGATAT	1800
QY	622	TyrlEugIyMcPhemEaSnSerPheAlaValPheGInGluCySgIyLeuTrpValIeu	641
Db	1801	TATTTAGGCAAGTTCAGTAAATCTTTTGCACTCTTCAGGAATGTGACCTGGGTATTG	1860
QY	642	ThraSpAlaAsnLeuThrLysAspTyrIleAspGlyValTyraSPasnaIaGluTyraIa	661
Db	1861	ACAGATGCAAACTTCACGAAGATTATATGATGGTGTATTATGCAATGACGAATATGCT	1920
QY	662	GluArpPheMcGluGluAsnGluGlyhIaIleValAspIlehIaSPheSerLeuGly	681
Db	1921	GAGAGGTTTATGAGGAAATGAAGGACATATTTGATGATTCAATGACTTATCTTTGGCT	1980
QY	682	SeSerProhIsvAlaArgLysHIpheProGluThrTrpIleTrpLeuAspThraSmeC	701
Db	1981	AGCAGTCCACATGTCGGAAGCATTTTCCAGACCTTGAGATTGACTGACCAACAACATG	2040
QY	702	GlyTyraArgIleTyrgInGluPheGluValThraValProAspSerIleThSerTrpVal	721
Db	2041	GGTTCACGAAATTTACAAAGATTTTGAAGTAACTGTACCTGATTTCTATCACTCTTGCGT	2100
QY	722	AlaThrgIyPheValIleSerGluAspLeuGlyLeuGlyLeuThrThrProValGlu	741
Db	2101	GCTACTGTTTGTGATCTCTGAGGACCTGGGTCTTGACATMAACAATCACTCCAGTGGAG	2160
QY	742	LeuGlnAlaPheGlnProPhePheIlePheLeuAsnLeuProTyrsSerValIleArgIy	761
Db	2161	CTCCAAGCTTCCACCACTTTTATTTTGTGATCTTCCCTCTCGTATACAGAGT	2220
QY	762	GluGluPheAlaLeuGluIleThrIlePheAsnTyrlEuyLysAspAlaThhGluValLys	781
Db	2221	GAAATATTGCTTTGGAATACATCATATTCAATTATTGAAAGATGCCACAGAGTTAAG	2280
QY	782	ValIleIleGluYsSerAspLysPheAspIleLeuMcThrSerSerGluIleAsnAla	801
Db	2281	GTAATCATTTGAGAAAGATGACAAATTTGATATTCTTAATGACCTTCAATGAATATAAGCC	2340
QY	802	ThrgIyHIsGInGInThrLeuLeuValProSerGluAspGlyAlaThraValLeuPhePro	821
Db	2341	ACAGGCCACCCGACGAGACCTCTGCTGCCAGTACGATGGGGCAACTGCTTTTCCC	2400
QY	822	IleArpProThrHIsleuGlyGluIleProIleThraValThraIleAsuSerProThraIa	841
Db	2401	ATCAGGCCAACACACTGGGAGAAATTCATATCACAGTCAACAGCTCTTTCACCCACGTCT	2460
QY	842	SeRaSPaIaValThrgInMcIleLeuValLysAlaGluGlyIleGluYsSerTySer	861
Db	2461	TCTGATGCTGCACCCAGATGATTTTACTAAGCTGAGAGAAATGAAATCATATTTCA	2520
QY	862	GlnSerIleleuLeuAspLeuThraSPasnaIrgLeuGlnSerThrLeuTyThrLeuSer	881
Db	2521	CAATCCATCTTATTAGACTTACAGCAATATGCTACAGAGTACCCTGAAAACTTTGAGT	2580
QY	882	PheSerPheProPheAsnThraValThhGlySerGluArgValGlnIleThraIaIleGly	901
Db	2581	TTCTCATTTCTCTCTTAATACAGTACCTGGCAGTAAAGATTTCAAGTACACGCAATTGGA	2640
QY	902	AspValLeuGlyProSerIleAsnGlyLeuAlaSerLeuIleAsgMcProTyrgIyCys	921
Db	2641	GATGTTCTTGTCCTTCCATCAATGGCTTACCTCATTTGATTCGAGATGCCCTTATGGCTGT	2700
QY	922	GlyGluGlnAsnMcIleAsnPheAlaProAsnIleTyrlleuAspTyrlEuyThrLys	941
Db	2701	GGTGAACAGACATGATTAATTTTGCTCAAAATATTTAACATTTTGGATTATCTGACATAA	2760
QY	942	LysIySgInLeuThraSPasnaIleuLysGluYsaIleAsuSerPheMcThrgInGlyTy	961
Db	2761	AAGAAACAACATGACAGATTAATTTGAAAGAAACCTTTTCAATTTATGAGGCAAGGTAC	2820

QY	962	GLNARGGLLEULEUENLYRGLNARGGLNLSPLYSERPHESERIALAPHEGLVANSITRASP	981
DB	2821	CAGAGAGAACTCTCTATCAGAGGAGAGATGGCTTTCACTGCTTTTGGGAAATTATGAC	2886
QY	982	PROSERGIYSERTHTTRPLEUSERIALAPHEVALLEUARGCYSPHELEUGLUALAASP	1001
DB	2881	CCTTCCTGGGACACTTGGTGTGCAGCTTTTGTCTTTTAAGATGTTTCTTGAAGCCGACTCT	2945
QY	1002	TYRILEASPIILEASPGINASNVALLLEUHSARGTHRTRYTHRTRPLEULYGLYHISGLN	1021
DB	2941	TACATAGATATATGATTCAGAAATGTTGTTCACAGAAACATPACCTGGCTTAAAGACATCAG	3000
QY	1022	LYSSEARANGLYGULPHETRPASPROGLVARGVALIIEHISSENGULEUGLGLYGLY	1041
DB	3001	AAATCCACGGGTGAATTTGGGATGCCAGAAAGAGTATCTATGAGACTTCAAGAGTGGC	3066
QY	1042	ASNLYSSERPROVALTHRLLEUTHALATYRILEVALTHRSERLEULEGLTYRARGHYS	1061
DB	3061	AATAAAGTCCAGTRACACTTACAGCCTATATTTGTACTTCTCTCGGATATRGAAG	3122
QY	1062	TYRGLNPROASNIIEASPVALLINGLUSERLIEHISPHLEUGLUSERGLUPHESERARG	1081
DB	3121	TATCAGCCTTAACATTTGATGTGCAAGAGCTATCATTTTGGAGTCCGAATTCAGTRAGA	3188
QY	1082	GLYLISERASPAANTYRTHRLEUALALEULEHETHRYRVALALEUSSESERVALGLYSER	1101
DB	3181	GGAAATTCAGCAATTTACTTACCTTACGCCCTTATATTAAGATTGATTCATCGTGGGGAGT	3240
QY	1102	PROLYSALALYSGLUALALEUASMERLEUTHTRTPARGALLAGLUNGGLYGLYMET	1121
DB	3241	CCTTAAGCGAAGGAAAGCTTTGAAATATGTGTGACTTGGAGAGCAGAACAAAGAGTGGCATG	3300
QY	1122	GLNPHETRPVALSERSEGIUSERLYLSLEUSERASPSERTTPGINPROARGSERLEUASP	1141
DB	3301	CAATTCGGGTGCATCAGAGTCCAAACTTTCTGACTCTGCGAGCCACGGCTCCCTGGAT	3366
QY	1142	ILEGLVALALALATYRALALEULEUSERHISPHLEUGLNPHEGLNTHRSERGLUGLY	1161
DB	3361	ATTGAAGTTCAGGCTTATGCACCTCTCCACACTTCTTACAAATTCAGACTTCGAGCGGA	3420
QY	1162	ILEPROILEMETARGTTRPLEUSERARGGLNARGHANSERLEUGLYGLYPHALSERTHR	1181
DB	3421	ATCCCAATTAAGAGGTGGCTAAGCAGGGAAGAAATAGCTTGGGGGTGTTTGCACTACT	3480
QY	1182	GLNAPSPTHTRVALALALEULYSALALEUSERGLUPHEALALALEMETASNTHRGLU	1201
DB	3481	CAGAGATNCACGTGGCTTTAAAGCTCTGTCTGAATTTGCAAGCCCTTATGAATACAGAA	3540
QY	1202	ARGTHRANILEGINVALTHRVALTHRGLYPROSESERPROSERPRO-----	1217
DB	3541	AGGACAAATATCCAAAGTGACCGTGAACGGGGCTTACGTCAACAAATCTCGTAAAGTTTCTG	3600
QY	1218	-----LEUALALVALGLNPROMET	1224
DB	3601	ATTGACACACAAACCGCTTACTCTTCAGACAGCAGCTTGTGTGGTACACCCATG	3660
QY	1225	ALALVALASNIILESERIALAENGLYPHEGLYPHEALALIECYSGINLEUASNVALLTYR	1244
DB	3661	GCACTTAATATTTCCGCAAAAGTGTITGGATTTCATTTCTCAGCTCAATCTGTATAT	3720
QY	1245	ASNVALLYBALASERGIYSERSEARARGHARGHRSERILEGINASNGLUALABHE	1266
DB	3721	AATGTGAAGGCTTCGGGTCTTTTACGAAGAAGAGATCTATCCAAAATCAGAGAAGCCTTT	3780
QY	1265	ASPLEUENSPVALALAVALLYSGLUALSNLYEASPAASPLEUENHISVALASPLEUASNVALL	1288
DB	3781	GATTATGATGTGTCTGTAAGAAAGAAATTAAGAAATATCTCATCATGTGGATTGAATGTG	3840
QY	1285	CYTHRHSERPHESERGIYPROGLVARGSERGIYMETALALEUMETGLIVALASNLEULEU	1304
DB	3841	TGTACAAAGCTTTTGGGCGCGGGAGTAGTGACGTGCTCTTATGGAAGTTAACCTATTA	3900
QY	1305	SERGLYPHEMETVALPROSERGIUALALIESEERLEUSERGIUNHTRVALLYSLYVALGLU	1324

Db 3901 AGTGGCTTATGAGCCCTTCAGAACATTTCTCTGAGCGAGACAGTGAAGAAAGTGA 3960  
 Qy 1325 TTTAAATGAGTATGAGTATGAGTATGAGTATGAGTATGAGTATGAGTATGAGTAT 1344  
 Db 3961 TATATGAGTATGAGTATGAGTATGAGTATGAGTATGAGTATGAGTATGAGTAT 4020  
 Qy 1345 AATATGAGTATGAGTATGAGTATGAGTATGAGTATGAGTATGAGTATGAGTAT 1364  
 Db 4021 AATATGAGTATGAGTATGAGTATGAGTATGAGTATGAGTATGAGTATGAGTAT 4080  
 Qy 1365 ValAspTyrTyrGluProArgArgGlnAlaValArgSerTyrAsnSerGluValLeu 1384  
 Db 4081 GTGATTACTATGAGCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4140  
 Qy 1385 SerSerCysAspLeuCysSerAspValGlnGlyCysArgProCysGluAspGlyVal 1403  
 Db 4141 TCTCTGCTGACCTTGGAGTGAATGCCAGGCGCTGCTGAGAGATGAGCT 4197  
 RESULT 14  
 ABQ79968 ID ABQ79968 standard; cDNA; 3535 BP.  
 AC ABQ79968;  
 DT 23-DEC-2002 (first entry)  
 XX 23-DEC-2002 (first entry)  
 DE Human CD109 K15 protein encoding cDNA.  
 XX CD109; alpha2 macroglobulin; thioester; cerebroprotective; cardiant;  
 KW immunosuppressive; haemostatic; anticoagulant; thrombolytic; human;  
 KW cardiovascular; vasotropic; gene therapy; CD109 K15; gene; ss.  
 XX Homo sapiens.  
 OS  
 XX  
 FH Key Location/Qualifiers  
 FH CDS 113..4450  
 FT /\*tag= a  
 FT /product= "CD109 K15"  
 XX  
 PN WO200270696-A2.  
 PD 12-SEP-2002.  
 XX  
 PF 07-MAR-2002; 2002WO-CA00292.  
 XX  
 PR 07-MAR-2001; 2001US-273814P.  
 PA (SCHU/) SCHUH A.  
 PA (SUTH/) SUTHERLAND R D.  
 XX  
 PI Schuh A, Sutherland RD;  
 DR WPI; 2002-713450/77.  
 DR P-PSDB; ABB82169.  
 XX  
 PT New CD109 nucleic acids and polypeptides, useful in gene therapy,  
 PT particularly for treating strokes, myocardial infarctions, thrombosis,  
 PT thrombocytopenia, autoimmune diseases, or organ or bone marrow  
 PT transplantation -  
 XX  
 PS Claim 1; Fig 4a; 156pp; English.  
 XX  
 CC The invention relates to isolated nucleic acid molecules encoding CD109  
 CC polypeptides. These nucleic acid molecules include the human cDNA  
 CC sequences comprising CD109 K1, CD109 K1-H7, CD109 K15 or their variants.  
 CC CD109 is a novel member of the alpha2 macroglobulin (alpha2M)/C3, C4, C5  
 CC family of thioester-containing proteins. The CD109 polypeptides can be  
 CC expressed by standard recombinant methodology. The CD109 nucleic acid,  
 CC CD109 polypeptide or its mimetic, or the anti-CD109 antibody is useful  
 CC for treating or detecting a disease or disorder, e.g. conditions  
 CC associated with endothelial activation, platelet activation, activation  
 CC of the coagulation or fibrinolytic systems, activation of T lymphocytes

CC and of the complement system, quantitative or qualitative abnormalities  
 CC of platelet function, increased or impaired platelet aggregation and  
 CC activation, increased or impaired activation of the coagulation and/or  
 CC fibrinolytic systems, or impaired or increased immune activation. These  
 CC are also useful for treating cardiovascular disorders, stroke, myocardial  
 CC infarction, thrombosis, embolism, peripheral vascular disease,  
 CC thrombocytopenia, thrombocythemia, autoimmune diseases, organ  
 CC transplantation, or bone marrow transplantation. The CD109 nucleic acids  
 CC or their antisense nucleotide sequence are useful in gene therapy for  
 CC treating these conditions. The present sequence represents the human  
 CC CD109 K15 cDNA sequence.  
 XX  
 SQ Sequence 3535 BP; 1108 A; 675 C; 731 G; 1021 T; 0 other;  
 Alignment Scores:  
 Pred. No.: 0 Length: 3535  
 Score: 5470.00 Matches: 1062  
 Percent Similarity: 100.00% Conservative: 1  
 Best Local Similarity: 99.91% Mismatches: 0  
 Query Match: 74.44% Indels: 0  
 DB: 24 Gaps: 0  
 US-10-020-095-4 (1-1428) x ABQ79968 (1-3535)  
 Qy 1 MetGlnGlyProProLeuLeuThrAlaAlaHisLeuLeuCysValGlyThrAlaAlaLeu 20  
 Db 113 ATGAGAGGCCACCGCTCTGACCGCGCCACCTCTGCTGAGTGAACCGCGGCTG 172  
 Qy 21 AlAlaValAlaProGlyProAlaGlyPheLeuValThrAlaProGlyLeuLeaProGlyGly 40  
 Db 173 GCGGTGCTCCCGGCTCGGTTTGTGTGACAGCCCGAGAGATCATCAGCCCGAGGA 232  
 Qy 41 AsnValThrIleGlyValGluLeuLeuGluHisCysProSerGlnValThrValAla 60  
 Db 233 AATGTACATTTGGGGTGGAGCTTCTGGACATCTGCCCTTCAAGCTGATGAAAGCG 292  
 Qy 61 GluLeuLeuLysThrAlaSerAsnLeuThrValSerValLeuGluAlaGluGlyValPhe 80  
 Db 293 GAGCTGCTCAAGACAGATCAAACTCATCTGCTGCTGAGAGCAGAAAGGCTTTT 352  
 Qy 81 GluLysGlySerPheLysThrLeuThrLeuProSerLeuProLeuLeuSerAlaAspGlu 100  
 Db 353 GAAAGAGCTCTTTTAAAGACACTTACTCTTCACTACCTCTGAGACAGTGCAGATGAG 412  
 Qy 101 IleTyrGluLeuArgValThrGlyArgThrGluAspGluIleLeuPheSerAsnSerThr 120  
 Db 413 ATTATGAGCTACGCTGTAACCGAGAGTACCCAGAGTGAAGATTTATTTCTTAATAGTACC 472  
 Qy 121 ArgLeuSerPheGluThrLysArgIleSerValPheIleGlnThrAspLysAlaLeuTyr 140  
 Db 473 CGCTTATCATTTGAGACCAAGAGATATCTGTTCAATCAACAGCAAGCCTATATAC 532  
 Qy 141 LysProLysGlnGluValLysPheArgIleValThrLeuPheSerAspPheLysProTyr 160  
 Db 533 AAGCCAAAGCAAGAGAGAGATTCGATTTGTACACTCTTCCAGATTTTAAAGCCTTAC 592  
 Qy 161 LysThrSerLeuAsnIleLeuIleLysAspProLysSerAsnLeuIleGlnIntProLeu 180  
 Db 593 AAAACCTCTTTAAACATTTCAATTAAAGACCCCAATCAATTTGATCCAAAGTGTG 652  
 Qy 181 SerGlnIleSerAspLeuGlyValIleSerLysThrPheGlnLeuSerSerHisProIle 200  
 Db 653 TCACAAAGATGATCTTGAGAGTCAATTCCAAAACTTTTCAAGTATCTTCCATCCATA 712  
 Qy 201 LeuGlyAspTyrPheSerIleGlnValAlaAsnAspGlnThrTyrTyrGlnSerPheGln 220  
 Db 713 CTGTGTAACGTGCTATTCATCAAGTCAAGTGAACAGACAGATATATCATCATTTTCAAG 772  
 Qy 221 ValSerGluTyrValLeuProLysPheGluValThrLeuGlnThrProLeuTyrCysSer 240  
 Db 773 GTTTCACAAATATGATTAACAAATTTGAAGTGAATTTGAGACACCATATATATGTTCT 832  
 Qy 241 MetAsnSerLysHisLeuAsnGlyThrIleThrAlaLysTyrThrTyrGlyLysProVal 260

```

Db      833 ATGAATTCAGAGCATTTAAATGATGACCATCAGCGCAAGATATCAATATGGGAAGCCAGAG 892
Qy      261 LysGIAspValThrLeuThrPheLeuProLeuSerPheTrpGlyLysIleVal 280
Db      893 AAAGAGAGAGTAAAGCTTACATTTTACCTTATCTTTTGGGGAAGAAGAAATATT 952
Qy      281 ThrLysThrPheLysIleLeuGlySerAlaAsnPheSerPheAsnAspGluMetLys 300
Db      953 ACAAAAACATTTAAGATTAATGATCTGCAAACTTCTTTTATGATGAAGATGA 1012
Qy      301 AsnValMetAspSerSerSerSerGlyLeuSerGluTyrLeuAspLeuSerSerProGlyPro 320
Db      1013 AATGATATGATTCCTCAATAGACCTTTGCAATACCTGATCTATCTTCCCTGACCA 1072
Qy      321 ValGluIleLeuThrThrValThrGluSerValThrGlyIleSerArgAsnValSerThr 340
Db      1073 GTRGAATTTTAAACCACTGACAGAAATCACTTACAGATTTTCAAGAAATGTAAACACT 1132
Qy      341 AsnValPhePheLysGlnHisAspTyrIleIleGluPheAspTyrThrThrValLeu 360
Db      1133 AATGTCTCTTCAAGCAACATGATTAATCATTTGAGTTTGTATTAATCTACTGTCTTG 1192
Qy      361 LysProSerLeuAsnPheThrAlaThrValLysValThrArgAlaAspGlyAsnGlnLeu 380
Db      1193 AAGCCATCTCTCAACTTCAACAGCCACTGAGAGTAACTGTGCTGATGCAACCACTG 1252
Qy      381 ThrLeuGluGluArgArgAsnValValIleThrValThrGlnArgAsnTyrThrGlu 400
Db      1253 ACTCTGGAAGAAAGAAATATATGATGATACATACAGTACAGACAGAAATCTATCTAG 1312
Qy      401 TyrTrpSerGlySerAsnSerGlyAsnGlnLysMetGluAlaValGlnLysIleAsnTyr 420
Db      1313 TACTGGAGCGGATCTTAACAGTGAATTCAGAAATCGAAGCTGTCAGAAATTAATTTAT 1372
Qy      421 ThrValProGlnSerGlyThrPheLysIleGluPheProIleLeuGluAspSerSerGlu 440
Db      1373 ACTGTCCCAAGTGAACCTTTTAAGATGTAATCCCAATCTCGAGAGATTCACATGAG 1432
Qy      441 LeuGlnLeuLysAlaThrPheLeuGlySerLysSerSerMetAlaValHisSerLeuPhe 460
Db      1433 CTACAGTTAAGGCTTAATTTCTGTGTAGTAAAGATGAGATGCACTCATATCTGTGTT 1492
Qy      461 LysSerProSerLysThrTyrIleGlnLeuLysThrArgAspGluAsnIleLysValGly 480
Db      1493 AAGTCTCCATGATGAACATACATCCAACTAAACAAAGATGAATAATTAAGTGGGA 1552
Qy      481 SerProPheGluLeuValValSerGlyAsnLysArgLeuLysGluLeuSerTyrMetVal 500
Db      1553 TCGCCTTTTGAAGTGTGTAGTGTGCAACAAACGATTGAGAGATTAGCTATATGTA 1612
Qy      501 ValSerArgGlyGlnLeuValAlaValGlyLysGlnAsnSerThrMetPheSerLeuThr 520
Db      1613 GATTCAGAGGAGACAGTGTGTGCTGTAAGAAACAAATTCACAAATGTTCTCTTAAACA 1672
Qy      521 ProGluAsnSerTyrThrProLysAlaCysValIleValTyrTyrIleGluAspAspGly 540
Db      1673 CCAAGAAATTTCTTGGACTCCAAAGCCTGTATATTGTATTAATTAAGATGAAGG 1732
Qy      541 GluIleIleSerAspValLeuLysIleProValGlnLeuValPheLysAsnLysIleLys 560
Db      1733 GAAATTAATAGATGATGTTCTTAAATTCCTGTTCACTGTTTAAATAATAGATTAAG 1792
Qy      561 LeuTyrTrpSerLysValLysAlaGluProSerGluLysValSerLeuArgIleSerVal 580
Db      1793 CTATATTTGAGTAAAGTGAAGCTGAACCATCTGAAGAAAGTCTCTTGTGAGATCTGTG 1852
Qy      581 ThrGlnProAspSerIleValGlyIleValAlaValAspLysSerValAsnLeuMetAsn 600
Db      1853 AACAGCCTGACCTCCATAGTGGGATTTGAGCTGTTGACAAAGTGTGAATCTGAAGAT 1912
Qy      601 AlaSerAsnAspIleThrMetGluAsnValValHisGluLeuGluLeuTyrAsnThrGly 620
|||||
Db      1913 GCCTTAATGATATTAACAATGGAATAATGTGCTCATGAGTGTGAACTTTATTAACAAGGA 1972
Qy      621 TyrTyrLeuGlyMetPheMetAsnSerPheAlaValPheGlnGlyCysGlyLeuTrpVal 640
Db      1973 TATTAATTTAGGATGTTCATGAATTTCTTTTGCAGTCTTTTCAGGAATGTGACTGTGGTA 2032
Qy      641 LeuThrAspAlaAsnLeuThrLysAspTyrIleAspGlyValTyrAspAsnAlaGluTyr 660
Db      2033 TTGACAGATGCAAACTCCAGAAAGATTAATATTGATGTGGTTTATGACAAATGAGAAATAT 2092
Qy      661 AlaGluArgPheMetGluGluLysAsnGlyLysIleValAspIleHisAspPheSerLeu 680
Db      2093 GCTGAAGGTTTATGAGAGAAATGAAGACATATTGATGATTAATCATGACTTTCCTTGG 2152
Qy      681 GlySerSerProHisValArgLysHisPheProGluThrTrpIleTrpLeuAspThrAsn 700
Db      2153 GGTAGACGTCCACATGTCGAAAGCATTTTCCAGAGACTTGAATTTGGCTAAGACACCAAC 2212
Qy      701 MetGlyTyrArgIleTyrGlnGluPheGluValThrValProAspSerIleThrSerTrp 720
Db      2213 ATGGGTACAGGATTTTACCAAGAAATTTGAAGTAACTGTAACCTGATCTATCACTTCTGG 2272
Qy      721 ValAlaThrGlyPheValIleSerGluAspLeuGlyLeuGlyLeuThrThrProVal 740
Db      2273 GTGGCTACTGTTTGTGTATCTCTGAGGACCTGGGCTTGAACCAACTACTCCAGTG 2332
Qy      741 GluLeuGlnAlaPheGluProPhePheIlePheAsnLeuProTyrSerValIleArg 760
Db      2333 GAGCTCCAGCCTTCCAACTTTTTCATTTTGTGAATCTTCCCTACTCTGTATTCACA 2392
Qy      761 GlyGluGluPheAlaLeuGluIleThrIlePheAsnTyrLeuLysAspAlaThrGluVal 780
Db      2393 GGTGAAGATTTGCTTTGGAATTAATCTAATTAATTAATTAATTAATTAATTAATTAAT 2452
Qy      781 LysValIleIleGluLysSerAspLysPheAspIleLeuMetThrSerSerGluLeuAsn 800
Db      2453 AAGGTAAATCATTTGAGAAAGTCAAAATTTGATTAATGATCACTTCAATCAATTAATTAAT 2512
Qy      801 AlaThrGlyHisGlnGluThrLeuLeuValProSerGluAspGlyValaThrValLeuPhe 820
Db      2513 GCACAGGCCACACACAGACCTTCTGTGTTCCCACTGAGGAGTGGGCAACTGTCTTTT 2572
Qy      821 ProIleArgProThrHisLeuGlyGluIleProIleThrValThrAlaLeuSerProThr 840
Db      2573 CCCATCAGGCCAACATCTGGAGAAATCTTAACAGATCAAGTCAAGCTTTTCAACCCACT 2632
Qy      841 AlaSerAspAlaValThrGlnMetIleLeuValLysAlaGluGlyIleGluLysSerTyr 860
Db      2633 GCTTCTGATGCTGTCAACCAAGATTTTATGTAAGGCTGAAGAGATGAATAATCATAT 2692
Qy      861 SerGlnSerIleLeuLeuAspLeuThrAspAsnAlaGluLeuGlnSerThrLeuLysThrLeu 880
Db      2693 TCACATCCATCTTATTAAGACTGACAGTACATAGCTACACAGTAACTGTAACAACTTGG 2752
Qy      881 SerPheSerPheProProAsnThrValThrGlySerGluArgValGlnIleThrAlaIle 900
Db      2753 AGTTTCTCTTTCTCTCTTAATACAGTACGTGCAAGTGAAGAGCTTCAGATCTGCAAT 2812
Qy      901 GlyAspValLeuGlyProSerIleAsnGlyLeuAlaSerLeuIleArgMetProTyrGly 920
Db      2813 GGAGATGTTCTTGTCCTTCATCAATAGGCTTACCTCATTTGATTCGATGCTTATGGC 2872
Qy      921 CysGlyGluGlnAsnMetIleAsnPheAlaProAsnIleTyrIleLeuAspTyrLeuThr 940
Db      2873 TGTGGTGAACACAAACATGATTAATTTGTCTCAAAATATTTAATTTGATTAATCTGACT 2932
Qy      941 LysLysLysGlnLeuThrAspAsnLeuLysGluLysAlaLeuSerPheMetArgGlnGly 960
Db      2933 AAAAGAAACACATGACATTAATTTGAAGAAAGAAAGCTTTTCAATTAATGAGCAAGT 2992
Qy      961 TyrGlnArgGluLeuLeuTyrGlnArgLysAspGlySerPheSerAlaPheGlyAsnTyr 980
Db      2993 TACCAAGAGAACTTCTCATACAGAGGAGATGCTCTTTCAGAGTCTTTTGGGAATTAAT 3052

```



QY 241 MetAsnSerIysHisLeuAsnGlyThrIleThrAlaIysThrThrGlyLysProVal 260  
 DB 833 ATGAATTTCAAGCATTTAAATGATACCATCAGCGCAAAATATACATATGGAAGCAGCG 892  
 QY 261 LysGlyAspValThrLeuThrPheLeuProLeuSerPheTrpGlyLysLysAsnIle 280  
 DB 893 AAAGGAGAGATAACGCTTACATTTTACCTTTATCCTTTGGGAAAGAAATATAT 952  
 QY 281 ThrLysThrPheLysIleAsnGlySerIleAsnPheSerPheAsnAspGluGluMetLys 300  
 DB 953 ACMAAAACATTTAAGATTAATGATCTGCAAACTTCTTTATATATATAAGAGAGAA 1012  
 QY 301 AsnValMetAspSerSerAsnGlyLeuSerGluThrLeuAspLeuSerProGlyPro 320  
 DB 1013 AATGTAATGGAATCTTCAATAGCACTTTCGAATACCTGATCTATCTTCCCTGGACCA 1072  
 QY 321 ValGluIleLeuThrThrValThrGluSerValThrGlyIleSerArgAsnValSerThr 340  
 DB 1073 GTAGAAATTTTAAACCAAGTGCAGATCAGATCAGTTACAGTTATTCAGAAATGTAAAGCACT 1132  
 QY 341 AsnValPhePheLysGlnHisAspThrIleIleGluPheAspThrThrThrValLeu 360  
 DB 1133 AATGTTCTTCAAGCAACATGATTAATCATATGATTTTGTATTAATCTACTGCTTG 1192  
 QY 361 LysProSerLeuAsnPheThrAlaThrValLysValThrArgAlaAspGlyAsnGluLeu 380  
 DB 1193 AAGCATCTCTCAACTTCAAGCACTGAGAGTAACCTGTCCTGATGACCAACCAACTG 1252  
 QY 381 ThrLeuGluGluArgArgAsnValValIleThrValThrGlnArgAsnThrThrGlu 400  
 DB 1253 ACTCTTGAAGAAAGAAATATATGATCATTAACAGTACAGAGAAATCATACTGAG 1312  
 QY 401 TyrTrpSerGlySerAsnSerGlyAsnGluLysMetGluAlaValGluLysIleAsnThr 420  
 DB 1313 TACTGAGGCGGANTCAACAGTGAATCAGAAATGAAAGCTGTCAGAAATTAATAT 1372  
 QY 421 ThrValProGlnSerGlyThrPheLysIleGluPheProIleLeuGluAspSerSerGlu 440  
 DB 1373 ACTGTCCTCCCAAGTGAACCTTTAAGATGTAATTCCTCAATCTGAGAGATTCCTAGTGG 1432  
 QY 441 LeuGlnLeuLysAlaThrPheLeuGlySerLysSerSerMetAlaValHisSerLeuPhe 460  
 DB 1433 CTACAGATTGAAGCCCTAATTCCTTGGTAGTAATAAGTACATGCACTCATGCTGTT 1492  
 QY 461 LysSerProSerLysThrThrIleGlnLeuLysThrArgAspGluAsnIleLysValGly 480  
 DB 1493 AAGTCTCCATGATGAACATACATCCAACTAAACAAAGATGAAATATTAAGGTGGGA 1552  
 QY 481 SerProPheGluLeuValValSerGlyAsnLysArgLeuLysGluLeuSerThrMetVal 500  
 DB 1553 TCGCCTTTGAGTGGTGGTGTAGTGCACAAACATTAAGAGGTTAAAGCTATATGATA 1612  
 QY 501 ValSerArgGlyGlnLeuValAlaValGlyLysGlnAsnSerThrMetPheSerLeuThr 520  
 DB 1613 GATATCAGGGGACAGTTGGTGGCTGTAGGAAACAAATTTCAACATGTTCTCTTTAACA 1672  
 QY 521 ProGluAsnSerThrProLysAlaCysValIleValIleValIleGluLysAspGly 540  
 DB 1673 CCAGAAATTTCTTGACCTCCAAAGCCTGTATATGTATATTAAGATGATGGG 1732  
 QY 541 GluIleIleSerAspValLeuLysIleProValGlnLeuValPheLysAsnLysIleLys 560  
 DB 1733 GAATTTATATAGTATTTCTTAATAATCTCTGTTCACTGTTTAAAAATATAGATTAAG 1792  
 QY 561 LeuThrTrpSerLysValLysAlaGluProSerGluLysValSerLeuArgIleSerVal 580  
 DB 1793 CTAATATGAGTAAGTAAGTAAGTGAACATCTGAAGAAAGTCTCTTGAATCTCTGG 1852  
 QY 581 ThrGlnProAspSerIleValGlyIleValAlaValAspLysSerValAsnLeuMetAsn 600  
 DB 1853 ACACAGCCTGACCTCATATGTTGGATTTAGCTGTGACAAAGTGTGATCTGATGAAAT 1912

QY 601 AlaSerAsnAspIleThrMetGluAsnValValHisGluLeuGluLeuThrThrGly 620  
 DB 1913 GCCTTAATGATATATTAACAATGGAATAATGGCCATGAGTTGGAACCTTATTAACAAGGA 1972  
 QY 621 TyrThrLeuGlyMetPheMetAsnSerPheAlaValPheGlnGluCysGlyLeuThrVal 640  
 DB 1973 TATATTTAGGACATGTTCAATAAATTTTTCAGCTTTCAGGATGGAATGGACTCGGGCTA 2032  
 QY 641 LeuThrAspAlaAsnLeuThrLysAspThrIleAspGlyValIleAspAsnAlaGluThr 660  
 DB 2033 TTGACAGATGCACCAACTACAGAGATTAATATGATGGTTTATGACAAATGCACAAAT 2092  
 QY 661 AlaGluArgPheMetGluGluAsnGluGlyHisIleValAspIleHisAspPheSerLeu 680  
 DB 2093 GCTGAGAGTTTATGAGAGAAATGAAGACATATTTGATGATATTCATGACTTTTCTTTG 2152  
 QY 681 GlySerSerProHisValArgLysHisPheProGluThrTrpIleTrpLeuAspThrAsn 700  
 DB 2153 GGTACCACTCCACATGTCGCAAGCATTTTCCAGAGACTTGGATTTGGCTAGACCAAC 2212  
 QY 701 MetGlyThrArgIleThrGlnGluPheGluValThrValProAspSerIleThrSerThr 720  
 DB 2213 ATGGTTCCAGGATTTACCAAGATTTGAAGTACCTGATCTTATATCACTTCTTGG 2272  
 QY 721 ValAlaThrGlyPheValIleSerGluAspLeuGlyLeuThrThrThrProVal 740  
 DB 2273 GTGGCTATCGGTTTGTATCTCTGAGGACCTGGGTTGAGTACATCAACTCACTCACTG 2332  
 QY 741 GluLeuGlnAlaPheGlnProPhePheIlePheLeuAsnLeuProThrSerValIleArg 760  
 DB 2333 GAGCTCCAAAGCCTTCCAAACCAATTTTTCATTTTGTGAATCTTCCCTACTCTGTTATCAGA 2392  
 QY 761 GlyGluGluPheAlaLeuGluIleThrIlePheAsnThrLeuLysAspAlaThrGluVal 780  
 DB 2393 GGTGAAGATTTGCTTGGAAATTAATATTAATTAATTAATTAATTAATTAATTAATTAAT 2452  
 QY 781 LysValIleIleIleGluLysSerAspLysPheAspIleLeuMetThrSerSerGluIleAsn 800  
 DB 2453 AAGGTATATCTTGAAGAAAGTGAACAAATTTGATATTCATTAAGCTTCAAAATGAATAAT 2512  
 QY 801 AlaThrGlyHisGlnGlnThrLeuLeuValProSerGluAspGlyAlaThrValLeuPhe 820  
 DB 2513 GCCACAGGCCACACAGCAACACCTTCTGGTCCAGTGAGATGGGGCAACTGTTCTTTT 2572  
 QY 821 ProIleArgProThrHisLeuGlyGluIleProIleThrValIleThrAlaLeuSerProThr 840  
 DB 2573 CCCATCAGGCCCAACATCTGGAGAAATTCCTATCAACAGTCAAGCTCTTTCACCACT 2632  
 QY 841 AlaSerAspAlaValThrGlnMetIleLeuValLysAlaGluGlyIleGluLysSerThr 860  
 DB 2633 GCTTCTGATGCTGTCACCAATGATTTTATGTAAGCTGAAGATTAAGAAATATCATAT 2692  
 QY 861 SerGlnSerIleLeuLeuAspLeuThrAspAsnArgLeuGlnSerThrLeuLysThrLeu 880  
 DB 2693 TCACATATCATCTTAATTAATGATCTGATCAGCAATAGGCTACAGAGTACCTGAAAACTTGG 2752  
 QY 881 SerPheSerPheProProAsnThrValThrGlySerGluArgValGlnIleThrAlaIle 900  
 DB 2753 AGTTTCTAATTTCTCTTAATTAACGTGACTGGAGTGAAGATTCACATCACTGCAAT 2812  
 QY 901 GlyAspValLeuGlyProSerIleAsnGlyLeuAlaSerLeuIleArgMetProThrGly 920  
 DB 2813 GGAAGATGTTCTTGGCTTCCATCAATATGCTTACCTCATGATTCGGAATGCCCTTATGGC 2872  
 QY 921 CysGlyGluGlnAsnMetIleAsnPheAlaProAsnIleThrIleLeuAspThrLeuThr 940  
 DB 2873 TGTGTGAACAGAACATATTAATTTCTCCAAATATTTACATTTTGGATTAATCTGACT 2932  
 QY 941 LysLysLysGlnLeuThrAspAsnLeuLysGluLysAlaLeuSerPheMetArgGlnGly 960  
 DB 2933 AAAAAGAAACAACTGACAGATTAATTTGAAGAAAAAGCTCTTTCAATTAATAGGCAAGGT 2992  
 QY 961 TyrGlnArgGluLeuLeuThrGlnArgGluAspGlySerPheSerAlaPheGlyAsnThr 980

```
Db      2993  TACGAGAGAACTTCTATCAGAGGAGATGGCTCTTCAGTGTCTTTGGGAATTAT 3052
QY      981   AspProSerGlySerThrTripleuSerAlaPheValIleuArgCysPheLeuGluAlaasp 1000
Db      3053  GACCTTCTGGGAGACACTTGGTGTGCACCTTTGTGTTAAGATGTTCTTGAAGCCGAT 3112
QY      1001  ProTyrIleaspIleaspGlnasnValIleuHisargThrTyrThrTripleuLysGlyHis 1020
Db      3113  CTTACATAGATATTGATCAGAAATGTGTACACAGAACATACACTTGGCTTAAAGGACAT 3172
QY      1021  GlnLysSerAspGlyGluPheThrAspProGlyArgValIleHisSerGluLeuGlnGly 1040
Db      3173  CAGAAATCCAAGGAGGAAATTTGGGATCCAGAGAGAGTGAATCATAGAGCTTCAAGT 3232
QY      1041  GlnAsnLysSerProValThrLeuThrAlaTyrIleValThrSerLeuLeuGlyTyrArg 1060
Db      3233  GGCATTAAGAATCCAGTAACACTTACAGCTTATATTGTAACCTTCTCTCTGGGATATGA 3292
QY      1061  LysTyrGln 1063
Db      3293  AAGTATCAG 3301
```

Search completed: January 17, 2004, 01:15:04  
Job time : 983 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Comphen Ltd.

OM protein - nucleic search, using frame\_plus.p2n model

Run on: January 16, 2004, 21:02:03 ; Search time 198 Seconds

(without alignments)  
3183.309 Million cell updates/sec

Title: US-10-020-095-4

Perfect score: 7348  
Sequence: 1 MGGPPLTAHLLCVCTAAL.....HSSVIFRCFKLYEMELWL 1428

Scoring table: BLOSUM62

Xgapop 10.0	Xgapext 0.5
Ygapop 10.0	Ygapext 0.5
Fgapop 6.0	Fgapext 7.0
Delop 6.0	Delext 7.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DBV=xlp  
-Q=/cg2\_1/uspro\_pool\_p/us10020095/runatc\_15012004\_163823\_18719/app\_query.fasta\_1.1607  
-DB=Issued\_Patents\_NA -QFMT=fastap -SUFFIX=rn1 -MINMATCH=0.1 -LOOPCL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=bloms62 -TRANS=human40.cdi  
-LIST=45 -DOCALLIGN=200 -THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000  
-USER=US10020095 @CGN 1.1 118 @runatc\_15012004\_163823\_18719 -NCPU=6 -ICPU3  
-NO MMAP -LARGEOUTERY -NEG\_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOC  
-DBV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREDS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Issued Patents NA:\*  
1: /cg2\_6/prodata/2/ina/5A.COMB.seq:\*  
2: /cg2\_6/prodata/2/ina/5B.COMB.seq:\*  
3: /cg2\_6/prodata/2/ina/6A.COMB.seq:\*  
4: /cg2\_6/prodata/2/ina/6B.COMB.seq:\*  
5: /cg2\_6/prodata/2/ina/ACTUS.COMB.seq:\*  
6: /cg2\_6/prodata/2/ina/backfiles1.seq:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1441	19.6	4577	4	US-09-241-606-1
2	1346.5	18.3	4079	4	US-09-016-434-1114
3	973.5	13.2	5211	1	US-08-447-411-1
4	931	12.7	5056	2	US-08-793-126-2
5	931	12.7	5056	3	US-09-132-271-2
6	931	12.7	5067	3	US-09-142-334-23
7	894	12.2	5948	2	US-08-662-227-1
8	894	12.2	5948	4	US-09-017-947-1
9	885	12.0	5924	1	US-08-447-411-44
10	768.5	10.5	4138	1	US-08-447-411-75
11	768.5	10.5	4138	2	US-08-662-227-33
12	768.5	10.5	4138	4	US-09-017-947-33

C	13	245	3.3	19431	4	US-09-221-017B-1090	Sequence 1090, Ap
	14	234	3.2	339	4	US-09-311-352B-1	Sequence 1, Appli
	15	215.5	2.9	750	4	US-09-241-606-3	Sequence 3, Appli
C	16	203.5	2.8	14066	4	US-09-601-198-56	Sequence 56, Appli
	17	185	2.5	549	4	US-09-311-352B-3	Sequence 3, Appli
	18	180	2.4	4026	4	US-09-252-991A-10214	Sequence 10214, A
	19	174.5	2.4	7101	1	US-08-480-604A-9	Sequence 9, Appli
	20	174.5	2.4	7101	2	US-08-405-496A-9	Sequence 9, Appli
	21	174.5	2.4	7101	3	US-08-915-136-9	Sequence 9, Appli
	22	174.5	2.4	7101	4	US-08-957-310-9	Sequence 9, Appli
	23	174.5	2.4	7101	4	US-10-011-366-9	Sequence 9, Appli
	24	171	2.3	9873	4	US-09-328-352-1360	Sequence 1360, Ap
	25	165	2.2	28473	4	US-08-961-527-83	Sequence 83, Appli
	26	164.5	2.2	4977	4	US-09-071-035-257	Sequence 257, App
	27	164.5	2.2	4977	4	US-09-071-035-261	Sequence 261, App
	28	164.5	2.2	4977	4	US-09-071-035-265	Sequence 265, App
	29	164.5	2.2	13993	4	US-09-220-132-20	Sequence 20, Appli
	30	163.5	2.2	14070	4	US-09-108-006C-2	Sequence 2, Appli
C	31	163	2.2	37948	3	US-09-251-645-11	Sequence 11, Appli
	32	161	2.2	3312	4	US-09-601-198-58	Sequence 58, Appli
	33	161	2.2	7286	4	US-08-793-273C-1	Sequence 1, Appli
	34	161	2.2	7286	5	PCT-US95-11684-1	Sequence 1, Appli
C	35	160	2.2	5629	4	US-09-453-702B-243	Sequence 243, App
C	36	157.5	2.1	15016	4	US-09-601-198-60	Sequence 60, Appli
C	37	156	2.1	3636	4	US-08-961-527-78	Sequence 78, Appli
	38	156	2.1	30549	4	US-09-134-001C-322	Sequence 322, App
	39	155.5	2.1	9510	4	US-09-453-702B-256	Sequence 256, App
	40	154	2.1	9171	1	US-08-038-682-5	Sequence 5, Appli
	41	154	2.1	9171	1	US-08-302-832-5	Sequence 5, Appli
	42	154	2.1	9171	1	US-08-530-198-5	Sequence 5, Appli
	43	154	2.1	9171	2	US-08-469-880-5	Sequence 5, Appli
	44	154	2.1	9171	2	US-08-728-470-5	Sequence 5, Appli
	45	154	2.1	9171	2	US-08-617-697-5	Sequence 5, Appli

#### ALIGNMENTS

RESULT 1  
US-09-241-606-1  
Sequence 1, Application US/09241606  
Patent No. 6472140  
GENERAL INFORMATION:  
APPLICANT: Tanzi, Rudolph E.  
APPLICANT: Kovacs, Dora  
APPLICANT: Saunders, Aleister J.  
TITLE OF INVENTION: Alpha-2-Macroglobulin Therapies and Drug Screening Methods for  
FILE REFERENCE: 0609.4460003  
CURRENT APPLICATION NUMBER: US/09/241,606  
CURRENT FILING DATE: 1999-02-02  
NUMBER OF SEQ ID NOS: 27  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 1  
LENGTH: 4577  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: sig\_peptide  
LOCATION: (44)..(112)  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (44)..(4468)  
FEATURE:  
NAME/KEY: mat\_peptide  
LOCATION: (113)..(4468)  
US-09-241-606-1  
Alignment Scores:  
Pred. No.: 1.49e-142  
Score: 1441.00  
Percent Similarity: 46.20%  
Best Local Similarity: 29.06%  
Query Match: 19.61%  
Length: 4577  
Matches: 451  
Conservative: 266  
Mismatches: 551  
Indels: 284

DB: 4 Gaps: 59  
 US-10-020-095-4 (1-1428) x US-09-241-606-1 (1-4577)  
 QY 12 LeuLeuCyValValThrAlaAlaLeuAlaValAlaProGlyProArgPheLeuValThr 31  
 Db |||||  
 Db 86 CTCCTGCTCTCTCCGCCACAGACGCCCTCAGCTCTCTGAAAAACCGCATATATGCTCTG 145  
 QY 32 AlaProGlyIleIleArgProGlyGlyValValThrIleGlyValGluLeuGluHis 51  
 Db |||||  
 Db 146 GTCCCTCCCTGCTCCACACT---GAGACCATGAGAGAGGCTGTGCTCTCTGAGCTAC 202  
 QY 52 CysProSerGlnValThrValLeuAlaGluLeuLeuValThrAlaSerAlaLeuThrVal 71  
 Db |||||  
 Db 203 CTGATGAGACAGACCTGTAAGTCTCTCTGAGTCTGTGAGGGGAAAAACAGAGCTCT 262  
 QY 72 ---SerValLeuGluAlaGluGly---ValPheGluValGlySerPheValThrLeuThr 89  
 Db |||||  
 Db 263 TTCACTGACTGTGAGCGGAGAAATGACGTAAGCTCTGCTGTGCTCTC----- 310  
 QY 90 LeuProSerLeuProLeuAlaSerAlaAspGlu-----IleTyGluLeuArgVal 106  
 Db |||||  
 Db 311 -----GCTGTCCCAAAGTCTTCAATCAATGAGAGTAATGTTCTTCACTGTCCAAAGT 364  
 QY 107 ThrGlyArgThrGlnAspGluIleLeuPheSerAlaSerThrArgLeuSerPheGluThr 126  
 Db |||||  
 Db 365 AAAGGACCAACCAAGAA-----TTAAGAAAGCGGACCAAGATGATGTAAAGAAC 415  
 QY 127 LysAlaGlySerValPheIleGlnThrAspIleValAlaLeuTyTrpProValGlnVal 146  
 Db |||||  
 Db 416 GAGACACTGTGCTTGTGTCCAGACAGACAAATCAATCAAAACGAGGACAGACAGTG 475  
 QY 147 LysPheArgIleValIleThrLeuPheSerAspPheValProTyTrpValThrSerLeuAla 166  
 Db |||||  
 Db 476 AAATTTGCTGTCTCTCATGATGAGAAACCTTCAACCCCTGATGATGATTCCTCACTA 535  
 QY 167 Leu---IleLysAspProLysSerAspLeuIleGlnGlnThrLeuSerGlnGlnSerAsp 185  
 Db |||||  
 Db 536 GTATATCAATTCAGATCCCAAGAAATCGCATCGACAAATGAGCAAGTTTCCATGTAGAG 595  
 QY 186 LeuGlyValIleSerLysThrPheGlnLeuSerSerHisAspIleLeuGlyAspTrpSer 205  
 Db |||||  
 Db 596 GGTGCTCCCAAGCAATTTCTTTTCCCTCTCATCGAGCCCTTCACAGGCTCTCTACAG 655  
 QY 206 IleGlnValGlnValAspAsp-----GlnThrTyTrpGlnSerPheGlnValSerGln 223  
 Db |||||  
 Db 656 GTGTGTGATCAGAAAGAAATCAGGTGAGAGACAGACAC---CCTTTCACCTGTGAGGAA 712  
 QY 224 TyrValLeuProLysPheGluValThrLeuGlnThrProLeuTyTrpCysSerMetAspSer 243  
 Db |||||  
 Db 713 TTGTCTTCCCAAGTTTGAAGTACAGATACAGTGCACAAAGATATATCACCATCTTGGAA 772  
 QY 244 LysHisIleLeuAlaGlyThrIleThrAlaAlaValTyTrpGlyValProValIleGlyAsp 263  
 Db |||||  
 Db 773 GAAAGATGATGATGATGTGTGCTTATACATATGAGAAAGCTGTCTCTGACAT 832  
 QY 264 ValIleThrLeuThrPheLeuProLeu----- 271  
 Db |||||  
 Db 833 GTGACTGTGACATTTGCAAGAAATATAGTACGCTTCCACCTGCCACGCTGAAGATTTCA 892  
 QY 272 -----SerPheTrpGlyLys 276  
 Db |||||  
 Db 893 CAGGCTTTCTGTGAGAAATTCAGTGCAGCTAAACAGCCATGAGCTGCTTATACAGAA 952  
 QY 277 LysLysAlaLeuIleThrLysThrPheLysIleAlaGlySerAlaAlaPheSerPheAspSer 296  
 Db |||||  
 Db 953 GTAATA-----ACCAAGCTCTCCAGCTGAAGAG-----AAAGAG 988  
 QY 297 GluGluMetLysAlaValMetCysSerSerAlaGlyLeuSerGluTyTrpLeuAspLeuSer 316  
 Db |||||  
 Db 989 TATTAATATGAAA-----CTTCACTAGAGGCCACAGATCCA 1024  
 QY 317 SerProGlyProValGluIleLeuThr-----ThrValThrGluSerVal 331

Db 1025 GAGAAAGAAACAGGTGTGGAATGACTGGAAGCAGTCCAGTGAATCAACAGAACCATTA 1084  
 QY 332 ThrGlyIleSerArgAlaValSerThrAlaValPhePheGlyGlnHisAspTyTrpIle 351  
 Db |||||  
 Db 1085 ACCAAATCTCA---TTTGGAAGTGTGACTCACACTTGTGACAGGGA-----ATT 1132  
 QY 352 GluPhePheAspTyTrpThrValLeuLysProSerLeuAlaPheThrAlaThrValLys 371  
 Db |||||  
 Db 1133 CCCTCTTTGG-----CAG 1147  
 QY 372 ValThrArgAlaAspGlyAlaGlnLeuThrLeuGluGluArgArgAlaValIle 391  
 Db |||||  
 Db 1148 GTGGCTCTAGTAAATGGGAAAGCGTCCCTATACA-----ATAAAGTCATA 1195  
 QY 392 ThrValThrGlnArgAlaValTyTrpGluTyTrpSerGlySerAlaSerGlyAlaGlnLys 411  
 Db |||||  
 Db 1196 TTCAATC---AGAGAAATGAGCAAACTATTAC-----TCAATGTACACAGGATGAG 1246  
 QY 412 MetGluAlaValGln---LysIleAlaValThr-----ValProGlnSerGlyThrPhe 428  
 Db |||||  
 Db 1247 CATGGCTGTGTACAGTTCTCTATCAACACACCAAGTTATGGGTACCTCTTACTGTT 1306  
 QY 429 LysIleGluPhePro-----IleLeuGluAspSer 438  
 Db |||||  
 Db 1307 AGGCTCAATTAACAAGATCGTAGTCCCTGTACGGCTACCAAGTGGGTGCAGAAAGAAC 1366  
 QY 439 SerGluLeuGlnLeuLysAlaTyTrpLeuGlySerLysSerSerMetAlaValHisSer 458  
 Db |||||  
 Db 1367 GAAGAGGCACATCAACTGCTTAT----- 1390  
 QY 459 LeuPheLysSerProSerLysThrTyTrpIleGlnLeuLysThrArgAspGluAlaIleLys 478  
 Db |||||  
 Db 1391 CTGTGTCTTCCCAAGAAAGACTTGTGTCCACTTGAAGCCATGTCTCATGAAATCAACC 1450  
 QY 479 ValGlySer-----ProPheGluLeuValValSerGlyAlaLys-----Arg 492  
 Db |||||  
 Db 1451 TGTGTGCATCTACAGACAGTCCAGGACATTAATATCTGAATGAGAGCAACCTGTGGGG 1510  
 QY 493 LeuLysGluLeuSer-----TyrMetValValSerArgGlyGlnLeuValAlaValGly 510  
 Db |||||  
 Db 1511 CTAAAGAACTCTCTCTTATATGATATATGCAAAAGGAGCATGTGCCAACTGGG 1570  
 QY 511 -----LysGlnAlaSerThr-----MetPheSerLeuThr 520  
 Db |||||  
 Db 1571 ACTCATGAGCTGTGTGAGAGCAGAAAGCATGAAGGCCATTTTTCATCTCAATCCCT 1630  
 QY 521 ProGluAlaSerTrpThrProLysAlaCysValIleValTyTrpIleGluAspAspGly 540  
 Db |||||  
 Db 1631 GTGAAGTCAGACATTTGCTCTGTGCTCGGTCTCATATCTATCTGTTTACCTACCGGG 1690  
 QY 541 GluIleIleSerAspValLeuLysIleProValGlnLeuValPheLysAlaValIleLys 560  
 Db |||||  
 Db 1691 GACGTATTTGGGATTTCTGCAAAATATGATTTGAAATTTGTGTGCCCAACAAAGTGAGT 1750  
 QY 561 LeuTyTrpSerLysValLysAlaGluProSerGluLysValSerLeuAlaIleSerVal 580  
 Db |||||  
 Db 1751 TTGAGCTTACGCCCATCAACAAAGTCTCCAGCTCCACAGGCCCATCGGAGTCAACAGG 1810  
 QY 581 ThrGlnProAspSerIleValGlyIleValAlaValAlaAspLysSerValAlaLeuMetAla 600  
 Db |||||  
 Db 1811 GCT---CTTCAGTCCGCTGTGCGCTCCGCTGTGTGACCAAAAGCTGTGCTCATGAAG 1867  
 QY 601 AlaSerAlaAspIleThrMetGluAlaValHisGluLeu---GluLeuTyTrpAlaThr 619  
 Db |||||  
 Db 1868 CTGATGCTGAGCTCTGCGGCTCTGCTTACCACTGTACCAAAAAGAGCTCACT 1927  
 QY 620 GlyTyTrp-----TyrLeuGlyMet 625  
 Db |||||  
 Db 1928 GGTCTTCCCTGGGCTTTGAATGACAGAGAGTGAAGATCTCATATCTCATATATGTC 1987  
 QY 626 PheMetAlaSer-----PheAla 631  
 Db |||||

Db 1988 TATATTAAATGCAATCATATCTCAATGATCAAGTCAAAATGATAAGATATGTACAGC 2047  
Qy 632 ValPheGlnGluCysGlyLeuThrValLeuThrAspAlaAsnLeuThrLysAspTyrIle 651  
Db 2048 TTCTTAGAGACATGGGCTTAAAGGATTCACCAACTCAAGAATTCGTAAACCCAAATG 2107  
Qy 652 AspGlyValTyrAspAlaGlu-----Tyr 660  
Db 2108 TGTCCACAGCTTCACAGTATGAAATGATGACCTGAAGGTCTACGTAGAGTTTAT 2167  
Qy 661 AlaGluArgPheMetGluGluAsnGluGlnHisIleValAspIleHisAspPheSerLeu 680  
Db 2168 GAGTCAGATGTAATGGAAAGAGCCATGCACGCTGTCATGTTAAAG----- 2218  
Qy 681 GlySerSerProHis-----ValArgLysHisPheProGluThrTyrIleTyrLeu 697  
Db 2219 -----CTTCACACGAGACACCGTACCAAACTACTCCCTGACATGATCTGGAT 2269  
Qy 698 AspThrAsnMetCylTyrArgIleTyrGlnGluPheGluValThrValProAspSerIle 717  
Db 2270 TTGGTGTGTAAACTCAGCAGGGGTGGCTGAGGTAGATACAGTACCTGCACACATC 2329  
Qy 718 ThrSerThrValAlaThrGlyPheValIleSerGluAspLeuGlyLeuGlyLeuThr 737  
Db 2330 ACCGAGTGAAGGACAGGAGCGCTTCGCTGTGAAGATCTGAGACTGTGATCTTCC 2389  
Qy 738 ThrProValGluLeuGlnAlaPheGlnProPhePheIlePheLeuAsnLeuProTyrSer 757  
Db 2390 ACT---GCCCTCTCCGAGCTTCAGCCCTTCTTTGTGAGCTTCAATGCTTACTCT 2446  
Qy 758 ValIleArgGlyGluGluPheAlaLeuGluIleThrIlePheAsnTyrLeuLysAspAla 777  
Db 2447 GGTATTCGTGAGAGGCTTCACACTCAAGCCAGCTCTTAACCTTCCCAATG 2506  
Qy 778 ThrGluValLysValIleIleGluLysSerAspLysPheAspIleLeuMetThrSerSer 797  
Db 2507 ATCCGGTCAGTGTGAGCTGGAAGGCTCTCCGCTCTTCCCTGCTCCACAGTGAAG 2566  
Qy 798 Glu-----IleAsnAlaThrGlyHisGlnGlnThrLeuLeuValProSer 812  
Db 2567 GAACAAAGCCCTCACTGCATCTGTGCAAAAGGAGGCA----- 2605  
Qy 813 GluAspGlyAlaThrValLeuPheProIleArgProThrHisLeuGlyLeuIleProIle 832  
Db 2606 -----ACTGTGCTCGGAGATGAAACCCCAAGTCATTAGGAATGTGAATTC 2653  
Qy 833 ThrValThrAlaLeuSer-----ProThrAla 841  
Db 2654 ACTGTGAGCGCAGAGCGCACTAGATCTCAAGAGCTGTGAGACTGAGGTGCTTCAGTT 2713  
Qy 842 -----SerAspAlaValThrGlnMetIleLeuValLysAlaGluGlyIle 856  
Db 2714 CCTGAACACGAAAGAAAGACACAGTCATCAAGCTCTGTGTGTTAACTGAAGACAT 2773  
Qy 857 GluLysSerTyrSerGlnSerIleLeuAspLeuThrAspAsnArgLeuGlnSerThr 876  
Db 2774 GAGAAAGAAACAACATTCACACTCCACTTGTCCATCAAGGTGATGAGTTCT----- 2827  
Qy 877 LeuLysThrLeuSerPheSerPheProProAsnThrValThrGlySerGluArgValGln 896  
Db 2828 ---GAAAGAAATTATCCCTGAACCTGCACCAAAATGTGTAGAAAGAAATCTCCCGAGTTCT 2884  
Qy 897 IleThrAlaIleGlyAspValLeuGlyProSerIleAsnGlyLeuAlaSerLeuIleArg 916  
Db 2885 GTCTCACTTTTGGAGACATATTAGGCTCTGCATCAAAACACACAAATCTTCTCCAG 2944  
Qy 917 MetProTyrGlyCysGlyGluGlnAsnMetIleAsnPheAlaProAsnIleTyrIleLeu 936  
Db 2945 ATCCCTTAATGCGCTGTGAGAGCAATATGCTCTTCTCTTAACATCTATGTACTG 3004  
Qy 937 AspTyrLeuThrTyrLysLysGlnLeuThrAspAsnLeuLysGluLysAlaLeuSerPhe 956  
Db 3005 GATTATCTAAATGAACACAGCAGCTTACTCCAGATCAAGTCAAGGCCATTGGCTAT 3064

Qy 957 MetArgGlnGlyTyrGlnArgIleLeuLeuTyrGlnArgIleAspGlySerPheSerAla 976  
Db 3065 CTCACAGCTGGTATCAGAGACAGTTGAACATAAACAATAATGATGGCTCTCAGACAGC 3124  
Qy 977 PheGly---AspTyrAspProSer---GlySerThrThrLeuSerAlaPheValLeuArg 994  
Db 3125 TTGGGGAGCGATATAGGAGAGAACCAAGGCAACAGCTGGCTCAGACCTTTGTGTGAAG 3184  
Qy 995 CysPheLeuGluAlaAspProTyrIleAspIleAspGlnAsnValLeuHisArgThrTyr 1014  
Db 3185 ACTTTGCCCAAGCTCAGACCTTCACTGATCGATGAAGCAGACATTAACCAACCCCTC 3244  
Qy 1015 ThrThrPheLysGlyHisGlnLysSerAsnGlyGluPheThrAspProGlyArgValIle 1034  
Db 3245 ATATGGCTCTCCCAAGGACAGAGACAAATGGCTTTCTCAGAGAGCTGGGCTCAGCTGTC 3304  
Qy 1035 HisSerGluLeuGlnGlyValAsnLysSerProValThrLeuThrAlaTyrIleValThr 1054  
Db 3305 AACCAATGCCATTAAGGAGAGATAGATGAAGTGAACCTTCGCTCCGCTATATCASCATC 3364  
Qy 1055 SerLeuLeuGlyTyrArgLysTyrGlnProAsnIleAspValGlnGlnSerIleHisPhe 1074  
Db 3365 GCCCTTCTGAGATTCCTCTCAGCTCACTCACTGCTGTGTCGCAATGCCCTGTTTGC 3424  
Qy 1075 LeuGluSer-----GluPheSerArgLysIleSerAspAsnTyrThr 1088  
Db 3425 CTGGAGTCAAGCTGGAACAGACACAAAGAGGAGCCATGGC---AGCCATGTATATACC 3481  
Qy 1089 LeuAlaLeuIleThrTyrAlaLeuSerSerValGlySer---ProLysAlaLysGluAla 1107  
Db 3482 AAGACATGCTGGCTGATCTTGTGCTGGCAGGATTAACAGACAGAGAGAAAGTA 3541  
Qy 1108 LeuAsnMetLeuThrThrArgAlaGluGlnGluGlyGlyMetGlnPheThr----- 1124  
Db 3542 CTCAGATCACTTAATAGAGAACTGTGAAGAAACAACTGTTCATTGGAGAGGCCCT 3601  
Qy 1125 ValSerSerGluSerTyrLysSerAspSerThrGlnProArg-----SerLeuAspIle 1142  
Db 3602 CAGAAACCAAGGACCAAGTGGGAGATTTTACAAACCCCAAGCTCTCTGCTAGAGTG 3661  
Qy 1143 GluValAlaAlaTyrAlaLeuLeuSerHisPheLeuGlnPheGln-----ThrSer 1159  
Db 3662 GAGATGACATCTTAATGTCTCTGCTGCT---TATCTCAGGGCCACAGCCCAACCTCG 3718  
Qy 1160 GluGlyIlePro-----IleMetArgThrLeuSerArgGlnArgAsnSerLeu 1175  
Db 3719 GAGGACCTGACCTTCGAACCAACATCGTGAAGTGAACAGACAGACAGATCCCGAG 3778  
Qy 1176 GylGlyPheAlaSerThrGlnAspThrThrValAlaLeuLysAlaLeuSerGluPheAla 1195  
Db 3779 GGCGGTTCCTCTCCACCCAGACACAGTGGGTGCTCTCCATGCTCTGTCCAAATATGCA 3838  
Qy 1196 AlaLeuMetAsnThrGluArgThrAsnIleGlnValThrValThrGlyProSerSerPro 1215  
Db 3839 GCCGCGACATTTACC---AGGACTGGGAAGGCTGCACAGGTGACATATCAAGCTTCAAGG 3895  
Qy 1216 SerProLeuAlaValValGln----- 1222  
Db 3896 ACATTTTCCAGCAAAATTCACAGTGAACAACAATCGCTGTACTCAGACAGGTCTCA 3955  
Qy 1223 -----PrometAlaValAsnIleSerAlaAsnGlyPheGlyPheAlaIleCys 1238  
Db 3956 TTGCCAGAGCTGCTGGGGAATACAGCATGAAGTGAAGAGAAAGATGTCTTACCTC 4015  
Qy 1239 GlnLeuAsnValValTyrAsnValLysAlaSerGlySerSerArgArgArgSerIle 1258  
Db 4016 CAGACTCTCTTGAATATCAATAT-----CTC 4042  
Qy 1259 GlnAsnGlnGluAlaPheAspLeuAspValAlaValLys-----GluAsnLysAsp 1275  
Db 4043 CCAGAAAGGAAGATGCCCTTGTGCTTAGAGTGCAGACTGCTGCTCAAACTGTGAT 4102

QY 1276 AspleuAen---HisValAspleuAenValCysThrSerPheSerGlyProGly----- 1292  
 Db 4103 GAACCCAAAGCCGACACAGCTTCCAAATCTCCCTAAGTGTACATACAGGAGCCGC 4162  
 QY 1293 ---ArgSerGlyMetAlaLeuMetGluValAsnLeuLeuSerGlyPheMetValProSer 1311  
 Db 4163 TCTGCTCTCAACATGCGATGCTGTATGTGAAGTGTCTCTGGCTTATTCCTCCGAG 4222  
 QY 1312 GluAlaIleSerLeu-----SerGluThrValIlyValIlyValIlyValIlyValIlyVal 1328  
 Db 4223 CCAACAGTGAATATCTTGAAGAATCTTAACATGATGACCCGACGAGATGACGACGAC 4282  
 QY 1339 LysLeuAenLeuIlyrLeuAenSerValAsnGluThrGlnPheCysValAsnIleProAla 1348  
 Db 4283 CATGCTTGATTTACCTTGATTAAGGTGTCAATACAGACATGACATGCTTCTTCACGGT 4342  
 QY 1349 ValArgAenPheIlyValSerAenThrGlnAspIleValSerIleValAspIlyr 1358  
 Db 4343 CTGCAAGATGTCCAGTGAAGATCTCAACACGACATGAGAAAGTATGATTAATAC 4402  
 QY 1369 GluProArgGlnAlaValArgSerIlyrAsnSer 1380  
 Db 4403 GAGACGATGAGCTTGCAATCGCTGATCAATGCT 4438

## RESULT 2

US-09-016-434-1174  
 Sequence 1174, Application US/09016434  
 Patent No. 6500938

## GENERAL INFORMATION:

APPLICANT: Janice Au-Young  
 APPLICANT: Jeffrey J. Seilhamer  
 TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF SIGNALING  
 NUMBER OF SEQUENCES: 1490  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: INCYTE PHARMACEUTICALS, INC.  
 STREET: 3174 PORTER DRIVE  
 CITY: PALO ALTO  
 STATE: CALIFORNIA  
 COUNTRY: USA  
 ZIP: 94304

COMPUTER READABLE FORM:  
 MEDIUM TYPE: floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/09/016,434

FILING DATE: HEREMITH  
 CLASSIFICATION:  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER:

FILING DATE:  
 CLASSIFICATION:  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Zeller, Karen J.  
 REGISTRATION NUMBER: 37,071  
 REFERENCE/DOCKET NUMBER: PA-0002 US  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (650) 855-0555  
 TELEFAX: (650) 845-4166

INFORMATION FOR SEQ ID NO: 1174:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 4079 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 IMMEDIATE SOURCE:  
 LIBRARY: GENBANK  
 CLONE: g177869  
 US-09-016-434-1174

## Alignment Scores:

Pred. No.: 1,46e-132 Length: 4079  
 Score: 1346.50 Matches: 419  
 Percent Similarity: 46.66% Conservative: 238  
 Best Local Similarity: 29.76% Mismatches: 488  
 Query Match: 18.32% Indels: 263  
 Gaps: 54

US-10-020-095-4 (1-1428) x US-09-016-434-1174 (1-4079)

QY 12 LeuLeuCysValAlaCysThrAlaIleValAlaValAlaProGluProArgPheLeuValThr 31  
 Db 86 CTCTGGTCTCTCTGCGCCACAGACGCTCAGTCTCTGGAAGAAACGCAATATGATGTTCTG 145  
 QY 32 AlaProGlyIleIleArgProGlyGlyAsnValThrIleGlyValAlaGluLeuLeuGlnHis 51  
 Db 146 GTCCCTCTCCCTGCTCCACACT---GAGACCATGAGAAAGGCTGTGCTTCCCTGAGACTAC 202  
 QY 52 CysProSerGlnValThrValIlyValAlaGluLeuLeuIlyrThrAlaSerAenLeuThrVal 71  
 Db 203 CTGAATGAGACAGTGTACTGTATAGTGTCTCTTGAAGTCTGTCAAGGAGAAACAGAGCTTC 262  
 QY 72 ---SerValLeuGluIlyValGluGly---ValPheGluIlySerPheIlyrThrLeuThr 89  
 Db 263 TTCACTGACTGAGAGGAGGAGAAATGACGTAATCTCACTGTGTGCGCTTC----- 310  
 QY 90 LeuProSerLeuProLeuAenSerAlaAspGlu-----IleTyrgIleuArgVal 106  
 Db 311 -----GCTGTCCCAAGTCTTATCATCAATGAGAGATATGTTCTCACTGTCCAAAGTG 364  
 QY 107 ThrGlyArgThrGlnAspGluIleLeuPheSerAenSerThrArgLeuSerPheGluThr 126  
 Db 365 AAAGACCAACCCCAAGAA-----TTAAGAACGCGACCACTGATGTTAAGAAC 415  
 QY 127 LysArgIleSerValPheIleGlnThrAspIlyValAlaLeuIlyrThrPheGluGlnIlyVal 146  
 Db 416 GAGACAGTGTGTCTTGTTCACAGACAAATCAATCAATCAACAGGCGACAGAGTG 475  
 QY 147 LysPheArgIleValThrLeuPheSerAenPheIlyrThrIlyrThrSerLeuAenIle 166  
 Db 476 AATTTGTTGTGTCTCTCCATGATGAATGAATCTTACCCCTGAAATGATGATTCACATA 535  
 QY 167 Leu---IleLysAspProLysSerAenLeuIleGlnIlyrThrLeuSerGlnIlyrAsp 185  
 Db 536 GTATACATTCAGATCCCAAGAAATGCGATGCGCAATGCGAGATTTCCAGTTAGAG 595  
 QY 186 LeuGluValIleSerIlyrThrPheGlnIlyrSerHisProIleLeuGlyAspIlyrSer 205  
 Db 596 GGTGGCTCAAGCAATTTCTTTTCCCTCTCATGAGCCCTTCCAGGCTCTTCAAG 655  
 QY 206 IleGlnValGlnValAenAenPhe-----GlnThrTyrgIlyrThrPheGlnValIlyr 223  
 Db 656 GTGTGTGTACAGAAATCAAGTGAAGAGACAGACAC---CCTTACCCGTGAGAGAA 712  
 QY 224 TyrValLeuProLysPheGluValThrLeuGlnThrProLeuTyrgIlyrThrPheGlnVal 243  
 Db 713 TTGTGTTCTCCCAAGTTTGAAGTAACAGTCAACCAAGATATCAATCTTGGAA 772  
 QY 244 LysHisLeuAenGluIlyrThrIleThrAlaLysTyrgThrTyrgIlyrValIlyrValIlyr 263  
 Db 773 GAAGAGATGAATATACAGTGTGTGCTTATACATATGAGAAAGCTGTGCTTGGACAT 832  
 QY 264 ValThrLeuThrPheLeuProLeu----- 271  
 Db 833 GTACCTGTAGCATTTGCAAGAAATATAGTACGCTTCCGATCCACGCTGAAGATTC 892  
 QY 272 -----SerPheTrpGlyLys 276  
 Db 893 CAGGCTTCTGTGAAGAAATTAAGTGAAGAGCTTAACAGCATGCTGCTTATCAAGAA 952  
 QY 277 LysLysAsnIleThrIlyrThrPheIlyrIleAenGlySerAlaAsnPheSerPheAenAen 296  
 Db 953 GTAAAA-----ACCAAGGCTTCCAGCTGAAGAG-----AAGAG 988

QY 297 GUGUmetLysAsnValMetAspSerSerAsnGlyLeuSerGlyUtyrLeuAspLeuSer 316  
 DB |||||  
 DB 989 TATGAATGAAA-----CTTCACTAGAGGCCCAAGATCA 1024  
 QY 317 SerProGlyProValGluLeuThr-----ThrValThrGluSerVal 331  
 DB |||||  
 DB 1025 GAAGAAGAACAGTGTGGATTTGACTGGAGGAGCTCCAGTGAATCAAGAAACATA 1084  
 QY 332 ThrGlyIleSerArgAsnValSerThrAsnValPhePheGlnHisAspTyrIleIle 351  
 DB |||||  
 DB 1085 ACCAAACTCTCA---TTGTGAAGAGGACTCACTTTCCAGAGGA-----ATT 1132  
 QY 352 GluPhePheAspTyrThrThrValLeuLysProSerLeuAsnPheThrAlaThrValLys 371  
 DB |||||  
 DB 1133 CCTCTCTTGGG-----CAG 1147  
 QY 372 ValThrArgAlaAspIleAsnGlnLeuThrLeuGluGluArgAsnAsnValIle 391  
 DB |||||  
 DB 1148 GTGGCCCTGATGATGGAAGGCGTCCCTATACCA-----ATAAAGTCATA 1195  
 QY 392 ThrValThrArgAsnTyrThrGlyUtyrPheSerArgSerGlyAsnGlnLys 411  
 DB |||||  
 DB 1196 TTCATC---AGAGGAATAGAGCAACTTAC-----TCCATCTCACAGGATGAG 1246  
 QY 412 MetGluAlaValGln---LysIleAsnTyrThr-----ValProGlnSerGlyThrPhe 428  
 DB |||||  
 DB 1247 CATGGCCTTGATCACTCTATCAACACCAAGCTTATGGTACTCTCTTACTGTT 1306  
 QY 429 LysIleGluPhePro-----LleuGluLysPheSer 438  
 DB |||||  
 DB 1307 AGGGTCATATTCACAGAGATGATGCTGCTTACCGCTTACAGTGGATGTCACAAAGAACAC 1366  
 QY 439 SerGluLeuGlnLeuLysAlaTyrPheLeuGlySerLysSerSerMetAlaValHisSer 458  
 DB |||||  
 DB 1367 GAAGAGCACATCACACTGCTTAT----- 1390  
 QY 459 LeuPheLysSerProSerLysThrTyrIleGlnLeuLysThrArgAspLysAsnIleLys 478  
 DB |||||  
 DB 1391 CTTGTGTTCTCCCAAGAGGCTTTCACCTTGAAGCCCATGCTCATGACTGACCC 1450  
 QY 479 ValGlySer-----ProPheGluLeuValValSerGlyAsnLys-----Arg 492  
 DB |||||  
 DB 1451 TGTGGCCATCTCAGACAGTCCAGGACATTAATCTGAATGAGAGGACCCCTGCTGGG 1510  
 QY 493 LeuLysGluLeuSer-----TyrMetValValSerArgGlyGlnLeuValAlaValGly 510  
 DB |||||  
 DB 1511 CTGAAGACCTCTCTTTTATTAATCTGATTAATGGAAGGAGGACTTTCGAACTGGG 1570  
 QY 511 -----LysGlnAsnSerThr-----MetPheSerLeuThr 520  
 DB |||||  
 DB 1571 ACTCATGACAGTCTGTGTAAGGAGGAACATGAAGGCCATTTTCCATCTCAATCCCT 1630  
 QY 521 ProGluAsnSerTyrThrProLysAlaCysValIleValTyrTyrIleGluAspArgGly 540  
 DB |||||  
 DB 1631 GTGAAGTCAGACATGCTCTGCTGCTGCTGCTCATCATGCTGCTTTTACTTACCGGG 1690  
 QY 541 GluIleIleSerAspValLeuLysIleProValGlnLeuValPheLysAsnLysIleLys 560  
 DB |||||  
 DB 1691 GACGATTTGGGGATTTGCAAAATATGATGTAATTTGCTGGCCCAACAGAGGAT 1750  
 QY 561 LeuTyrTyrSerLysValLysAlaGluProSerGluLysValSerLeuArgIleSerVal 580  
 DB |||||  
 DB 1751 TTGAGCTTAGGCCATTCACAAAGTCTCCAGGCTCACAGCCCACTGGATGATCAGG 1810  
 QY 581 ThrGlnProAspSerIleValGlyIleValAlaValAspLysSerValAsnLeuMetAsn 600  
 DB |||||  
 DB 1811 GCT---CCTCAGCTCCCTGCGCCCTCCGCTGTGAGCAAAAGCTGCTGTCTATGAAG 1867  
 QY 601 AlaSerAsnAspIleThrMetGluAsnValValHisGluLeu---GluLeuTyrAsnThr 619  
 DB |||||  
 DB 1868 CCTGATGCTGAGCTCTGGCGCTCTCTGGTTTACAACTGCTACCAAAAGAGCTCTACT 1927  
 QY 620 GlyTyr-----TyrLeuGlyMet 625

DB 1928 GAGTTCCCTGGGCTTTGATATGACGAGGACGATGAAGACTGCATCAATCTCATATATGTC 1987  
 QY 626 PheMetAsnSer-----PheAla 631  
 DB |||||  
 DB 1988 TATATTAAATGGAATCATAATATCTCCAGTATGATCAATGAAGAAAGATATGAGCC 2047  
 QY 632 ValPheGlnGluCysGlyLeuThrValLeuThrAspAlaAsnLeuThrLysAspTyrIle 651  
 DB |||||  
 DB 2048 TTCCTAGAGACATGGGCTTTAAAGGCATTCACAACTCAAAAGATTCGTAACCCAAATG 2107  
 QY 652 AspGlyValTyrAspAsnAlaGlu-----Tyr 660  
 DB |||||  
 DB 2108 TGTCCACAGCTTCAACATATGAATGAATGATGAGCTGAAAGCTTACGTGAGTTTAT 2167  
 QY 661 AlaGluArgPheMetGluLysAsnGluLysIleValAspIleHisAspPheSerLeu 680  
 DB |||||  
 DB 2168 GAGTCAGATGTAATGGAGAGGCGCATGACGCTGGTGCATGTTGAAGG----- 2218  
 QY 681 GlySerSerProHis-----ValArgLysHisPheProGluThrTyrIleThrLeu 697  
 DB |||||  
 DB 2219 -----CCTCACAGGAGACCGGTACGAAGATCTTCCCTGAGACATGATCTGGAT 2269  
 QY 698 AspThrAsnMetGlyTyrArgIleTyrGlnGluPheGluValThrValProAspSerIle 717  
 DB |||||  
 DB 2270 TTGGTGTGTGTAATCTCAGAGGAGGCTGAGGATGAGTAAAGTCCGACACATC 2329  
 QY 718 ThrSerTyrValAlaThrGlyPheValIleSerGluAspLeuGlyLeuThrThr 737  
 DB |||||  
 DB 2330 ACCGATGGAAGGACGGGCGCTTGTGCTGTGAAGAGTGTGACTTGTATCTTCTCC 2389  
 QY 738 ThrProValGluLeuGlnAlaPheGlnProPhePheIlePheLeuAsnLeuProTyrSer 757  
 DB |||||  
 DB 2390 ACT---GCTCTCTCCAGCCTTCCAGCCTTCTTTTGAGAGCTTACAAATGCTTACTCT 2446  
 QY 758 ValIleArgGlyGluGluPheAlaLeuGluIleThrIlePheAsnTyrLeuLysAspAla 777  
 DB |||||  
 DB 2447 GTGATTCGTGAGAGGCTTACACTCAAGCCACAGGCTTAACTACTTCCCAATGTC 2506  
 QY 778 ThrGluValLysValIleIleGluLysSerAspLysPheAspIleLeuMetThrSer 797  
 DB |||||  
 DB 2507 ATCCGGGTCACTGTGCACCTGAAGCCTTCCGCTTCTGTCGCTGTCGCTGAGAGAG 2566  
 QY 798 Glu-----IleAsnAlaThrGlnIleGlnIleThrLeuValProSer 812  
 DB |||||  
 DB 2567 GAACAGGCGCTCAGTCACTGTGCAACGGCGGCA----- 2605  
 QY 813 GluAspGlyAlaThrValLeuPheProIleArgProThrHisLeuGlyGluIleProIle 832  
 DB |||||  
 DB 2606 -----ACTGTCTGGGAGTAAACCCCAAGTCAATTAAGAAATGTGAATTTC 2653  
 QY 833 ThrValThrAlaLeuSer-----ProThrAla 841  
 DB |||||  
 DB 2654 ACTGTGACGAGAGGACCTAGAGTCTCAAGCTGTGTGAGACTGAGCTCTTCACTT 2713  
 QY 842 -----SerAspAlaValThrGlnMetIleLeuValLysAlaGluGlyIle 856  
 DB |||||  
 DB 2714 CCTGAACACGGAAGAAAGACAGTCAAGGCTCTGTGTGTAACCTGAAGAGACTA 2773  
 QY 857 GluLysSerTyrSerGlnSerIleLeuLeuAspLeuThrAspAsnArgLeuGlnSerThr 876  
 DB |||||  
 DB 2774 GAGAAAGAAACAACTTAACCTCCCTACTTTGTGCATGAGTGGAGGTTCT----- 2827  
 QY 877 LeuLysThrLeuSerPheSerPheProProAsnThrValIleThrGlySerGluArgValGln 896  
 DB |||||  
 DB 2828 ---GAAAGATTAATCCCTGAATCTGCACCAATGTGTGAAGATGCGCCGACTTCT 2884  
 QY 897 IleThrAlaIleGlyAspValLeuGlyProSerIleAsnGlyLeuAlaSerLeuIleArg 916  
 DB |||||  
 DB 2885 GCTCAAGTTTGGAGACATATTAAGCTCTGCTCCATGCAAAACACACAAATCTTCTCAG 2944  
 QY 917 MetProTyrGlyCysGlyGluGlnAsnMetIleAsnPheAlaProAsnIleTyrIleLeu 936  
 DB |||||

Db	2945	ATGCCCTATGGCTGGAGAGACAAATATGGCTCTTGGCTCTTAACATCATATGACTG	3004
Qy	937	AerPTyLeuThrIyValyLeuGlnLeuThrIAspAsnLeuIyGluIyAlaLeuSerThe	956
Db	3005	GATTATCTAAATGAAAGACAGAGGCTTCTCCAGAGCAAGTCCAAAGGCCATGGCAT	3064
Qy	957	MetArgGlnGIyTYrGIyArgGluLeuLeuTYrGlnAspGIySerPheSerIa	976
Db	3065	CTCAACACTGGTTACCAAGACAGTGTAACTACAAACACTATGATGGCTCTTACAGACC	3124
Qy	977	PheGIy--AsnTYrAspProSer---GlySerThrTPLeuSerIaPheValIeuArg	994
Db	3125	TTTGGGGAGCGCATATATGACGAGAACAGGGCAACACTGGCTTCACAGCCCTTGTCTGAAG	3184
Qy	995	CysPheLeuGluAlaAspProTYrIleAspIleAspGlnAsnValIeuHiIaArgThTYr	1014
Db	3185	ACTTTTGGCCCAAGCTCGAGCCCTACCTCTTCATCGATGGAAGCAACATTAACCCAGCCCTC	3244
Qy	1015	ThrTPLeuIyGluYHISGlnIySerAsnGIyGluPheTPAspProGIyAlaValIle	1034
Db	3245	ATATGGCTCTCCCAAGGCGAAGAGCAATGGCTTTTCAGAGGCTCTGGGTCATGCTC	3304
Qy	1035	HisSerGIuLeuGlnGIyGIyAsnIySerProValThrLeuThrAlaTYrIleValIthr	1054
Db	3305	AACAATGCCATPAAGGAGGAGATGAAGAATGAAGAACCTCTCCGCTATATACACATC	3364
Qy	1055	SerIeuLeuGIyTYrArgIySerTYrGlnProAsnIleAspValGlnIuSerIleHisPhe	1074
Db	3365	GCCCTTCTGGAGATTCCTCTGCACAGTCACTCACTGTGTTCGGCAAGCCCTGTTTGC	3424
Qy	1075	LeuGIuSer-----GluPheSerArgGIyIleSerAspAsnThTYrThr	1088
Db	3425	CTGGAGTAGCCCTGGAGACAGCACAAAGAGGAGCCATGGC--AGCCATGTATATACC	3481
Qy	1089	LeuAlaIeuIleThTYrAlaLeuSerSerValGIySer---ProIyAlaIyIleGIuAla	1107
Db	3482	AAAGCACTGCTGGCCTATGCTTTTCCCTCGGAGGATTAACAGAGAAAGAGAAAGTA	3541
Qy	1108	LeuAsnMetLeuThTYrTPArgIaGIuGlnGIyGlyMetGlnPheTP-----	1124
Db	3542	CTCAAGTCACTTAATGAGGAAGCTGTGAAGAAAGACAACCTGTTCATTTGGAGGCGCCT	3601
Qy	1125	ValSerSerGIuSerIyLeuSerIAspSerTPGlnProArg-----SerIeuAspIle	1142
Db	3602	CAGAAACCCAAAGGCACAGAGGGGCAATTTTACGAACCCACAGGCTCCTCTGTGAGGTG	3661
Qy	1143	GluValAlaAlaIyAlaIyAlaLeuLeuSerHisPheLeuGlnPheGln-----ThSer	1159
Db	3662	GAGATGAATCTCTATGTGTCTCTGCT---TATCTCAGGGCCAGCCAGCCCAACCTCG	3718
Qy	1160	GluGlyIlePro-----IleMetArgTPLeuSerArgGlnArgAsnSerIeu	1175
Db	3719	GAGGACCTGACCTCTGCAACCAACATCGTGAAGTGTGATCAAGAGCACAAATGCCAG	3778
Qy	1176	GlyGIyPheAlaSerThrGlnAspPheThTYrValAlaIeuIySAlaLeuSerGIuPheAla	1195
Db	3779	GGCGGTTTCTCTCCACCCAGAGACACATGGTGCTCTTCATGCTCTGTCCAAATATGGA	3838
Qy	1196	AlaIeuMetAsnThGIuArgThTYrAsnIleGlnValThrValThrGIyProSerSerPro	1215
Db	3839	GCCCCACATTTACC--AGGACTGGGAAGGCTGCACAGGTGACTATCAAGTTTCAGGG	3895
Qy	1216	SerProIeuAlaValIaGln-----	1222
Db	3896	ACATTTTCCAGCAATTCACAGTGACAAACAACATGCCCTGTATCTGACAGAGCTCA	3955
Qy	1223	-----PrometAlaValAsnIleSerAlaAsnGIyPheGIyPheAlaIleCys	1238
Db	3956	TTGGCAGAGCTGCGCGGGAATACAGCATGAAAGTGAAGAGAGAAGATGTGTCTACCTC	4015
Qy	1239	GlnIeuAsnValValTYrAsnVal	1246
Db	4016	CAGACCTCTTGAATTAACAATATT	4039

```

RESULT 3
US-08-447-411-1
: Sequence 1, Application US/08447411
: Patent No. 5773243
: GENERAL INFORMATION:
: APPLICANT: FRITZINGER, DAVID C.
: APPLICANT: BREDEHORST, REINHARD
: APPLICANT: VOGEL, CARL-WILHELM
: TITLE OF INVENTION: DNA ENCODING COBRA C3, CVP1, AND CVP2
: NUMBER OF SEQUENCES: 81
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
: ADDRESSEE: P.C.
: STREET: 1755 S. Jefferson Davis Highway, Suite 400
: CITY: Arlington
: STATE: Virginia
: COUNTRY: U.S.A.
: ZIP: 22202
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patentin Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/447,411
: FILING DATE:
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/043,747
: FILING DATE: 07-APR-1993
: ATTORNEY/AGENT INFORMATION:
: NAME: Obion, No. 5773243man F.
: REGISTRATION NUMBER: 24,618
: REFERENCE/DOCKET NUMBER: 1126-101-0
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (703) 413-3000
: TELEFAX: (703) 413-2220
: TELEX: 248855 OPAT UR
: INFORMATION FOR SEQ ID NO: 1:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 5211 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: unknown
: TOPOLOGY: unknown
: MOLECULE TYPE: DNA (genomic)
: FEATURE:
: NAME/KEY: CDS
: LOCATION: 9...4961
: US-08-447-411-1

Alignment Scores:
Pred. No.: 1.37e-92 Length: 5211
Score: 973.50 Matches: 392
Percent Similarity: 40.564 Conservative: 280
Best Local Similarity: 23.656 Mismatches: 650
Query Match: 13.254 Indels: 335
DB: 1 Gaps: 62

US-10-020-095-4 (1-1428) x US-08-447-411-1 (1-5211)
QY 1 MetGInGlyProProLeuIeuThraAlaIaHisIeuIeuCyValCyThraAlaIaIeu 20
Db |:::| |::| |::| |::| |::| |::| |::| |::| |::| |::| |::| |::|
QY 9 ATGAGAGGATGGCTCTCATCTGTGCTCTCTATTGATGGTTTCCAGAGTCT--- 65
Db |::| |::| |::| |::| |::| |::| |::| |::| |::| |::| |::| |::|
QY 21 AlaValAlaProGlyProArgPheIeuValThraIaProGlyIleIleArgProGly 40
Db |::| |::| |::| |::| |::| |::| |::| |::| |::| |::| |::| |::|
QY 66 -----TCCACAGGGGCTCTCATACCTCATCAACCCCTGCTGTTTGGCAGACACA 119
Db |::| |::| |::| |::| |::| |::| |::| |::| |::| |::| |::| |::|
QY 41 AenValThriIleGlyValGIuIeuIeu---GIuHisCySProsArgIuValThriVal 59
Db |::| |::| |::| |::| |::| |::| |::| |::| |::| |::| |::| |::|
QY 120 GAAGACAAATTTGGTGGAGGCCCATGAGACACGATCTCCAAATGCGCTTGACATCTTT 179
Db |::| |::| |::| |::| |::| |::| |::| |::| |::| |::| |::| |::|

```



OY	60	ala-----gluleuleuysrthralaserAenleuthValSerValleuglu	75
Db	180	gTTCATGATTTTCCACGGAAGCAGAAACCTTGTTCCAAAGACAGATATATGATGACG	239
OY	76	AlaIugluYValPhegluldesglYserPheIysrthleuthrleuProserleuProleu	95
Db	240	CGAGAACAACAGTTGTGACT-----CCACTATTAAGGTTCTGCAAAAGACTG	290
OY	96	AenSerAlaAspGluIleTyrGluLeuArgValThrGlyArgThrGlnAspGluIleLeu	115
Db	291	AATAAGACTCCAAACAAATCAGTATGCTGGTTGGAAAGTAACGTGCTCCAGTGGCA	350
OY	116	PheSerAenSerThrArgLeuSerPheglunthrIlysaGIlleSerValPheIleGlnThr	135
Db	351	TTGGAAAAGGTGTTCTCTCTTCTTACCGAGATGGC-----TTGTGTTCATCCAGACA	404
OY	136	AspIlysaIaleuTyrIlyseProIysGlnGluValIlysaPheArgIlleValThrIleuPheSer	155
Db	405	GATTAAGGCACTTATACACCAAGGCTCTCCAGGCGGTATACGTGCTTCTTGCG-----	458
OY	156	AspPheIysProTyrIlyseThrSerLeuAsnIleLeuIleIysAspProIysSerAenLeu	175
Db	459	GATCAACAATGACACAGATGGACAAAACGTGATGTGACG-----	500
OY	176	IleGlnGlnThrPleuSerGlnGlnSerAspLeuGlyValIleSerIlyseThrPheGlnLeu	195
Db	501	-----TTTCAGACTCCAGAAAGCATTTGT-----GTC	527
OY	196	SerSerHleProIle-----	200
Db	528	AGTTTAAACCAAGTAAATTCATCAGGCTCGATCCGACCTTACAAATTATACCAAGCTTGTC	587
OY	201	--LeuGlyAspTrpSerIleGlnValGlnValAen-----AspGlnThrTyrTyr	216
Db	588	AGTTTGGGACATGGAAGCGCTGGCCAAATATGAACATTCACCAAGAAAGACTACACT	647
OY	217	GlnSerPheGlnValSerGluTyrValLeuProIysPheGlnValThrIleGlnThrPro	236
Db	648	GCATATTTTGATGTGACAGAAATATGTGTTACCAACTTTGAAGTCCGCTGCACCAACATCA	707
OY	237	-----LeuTyrCysSerMetAenSerIlyshIleuAenGlyThrIleThrAlaIys	253
Db	708	GATTAAGTTCTTATCAATGATGATGGGAT---AAAAATTCACAGTGTCTATCACTGCAAG	764
OY	254	TyrThrTyrGlyIlyseProValIlyseGlyAspValIThrLeuthrPheIleuProIysPhe	273
Db	765	TACTTATATGAAAGAAAGCTGGAAGT---GTGGCCTTGTGTCTTGGAGTCAAAATA	821
OY	274	TrpGlyIlyseIlyseAsnIle-----ThrIlyseThrPheIlyseIleAenGlySer	289
Db	822	GATGATGCTTAAAAAGATATTCACAGCTCATCTCAGACAGAAATTCAGATTATGATGAGAT	881
OY	290	AlaAenPheSerPheAsnAspGluGluMetIlysaAenValMetAspSerSerAenGlyLeu	309
Db	882	GGGGAAGCAACCTAAAAAGATACACATCACTTCCGATTTCAAGATCTCAATCAGCTT	941
OY	310	SerGluTyrLeuAspLeuSerSerProGlyProValGluIleLeuThrThrValProGlu	329
Db	942	GTTGGTCACTACTCTGTATGATCT-----GTACACGTGATTAACAGAA	983
OY	330	SerValThrGlyIleSerArgAsnValSerThrAenValPhePheIlyseGlnIleAspTyr	349
Db	984	TCAGGCACTGATATGTGATGACTGACGAAGCGGATCATATGTGACATCTCCAT	104
OY	350	IleIleGluPhePheAspTyrThrThrValIleuIlyseProSerLeuAenPheThrAlaThr	369
Db	1044	CAGACTCTACTTCACAAAAACCCCAAAATATTTCAAGCCGAGATGCAATATGAACCTGAG	1103
OY	370	ValIlysaValIThrArgAlaAspGlyAenGlnLeuThrIleuGluGluArgArgAenVal	389
Db	1104	GTGATATGTTACCAACCTGATGCTCACCAAGCTGCC-----CAT	1144
OY	390	ValIleThrValThrGlnArgAenTyrThrGluTyrTrpSerGlySerAenSerGlyAen	409

Db 1143 GTGCCAGTGTATACAGAGGCCATTTCATTCTAG-----GGAAACC 1181

Gly 410 GlnlysmetGuaIaValaGlnlyseIleasnYrThrValPro-----GlnserGly 426

Db 1182 ACTTTGAGTATAGGAGCTGCTAACCTACTTCTGAAACACACCACTGACATTCMAAGC--- 1233

Gly 427 ThrPheIysIleGlnPheProIleuGlnuAspSerSerGluLeuGlnIleuLysAlaYr 446

Db 1239 -----CTACCGATCATCTGTAGAACTAACACTGAGAGACTCCCAAGAGAA 1283

Gly 447 PheIeuGlySerIysSerSerMetAlaValHiserIleuPheIysSerProSerIysThr 466

Db 1284 CGCCAGGCAATAAAGTCCATGACAGCCACAGCCCTACAAACCAGGAGGGGTCTGAAAAC 1343

Gly 467 TyrIleGlnIleuLysThrArgAspGluAsnIleLysValGlySer-----ProheGlu 484

Db 1344 TATCTTCATGATGAGCCATTACATCTACAGAGATTAAGCCCCGAGATAACTTACCTGTCAAT 1403

Gly 485 LeuValValSerGlyAsn-----LysArgIeuLysGluLeuSerTyrMetVal 500

Db 1404 TTCAATGTGAGAGGGCAATGCAAAATTCATCGAACCAGATCAAAATTTTCACATACCTCTAA 1465

Gly 501 ValSerArgGlyGlnLeuValAlaValaGlyLys-----GlnAsnSer 514

Db 1464 TTGATATAAGGAAAGATTTCMAAGTTGGCGCAAGCCAGAGAGATGGCAGAGATCTG 1522

Gly 515 ThrMetPheSerIeuThrProGluAsnSerIyrPheProLysAlaCysValIleValYr 534

Db 1524 GTGACCATGAAATCTGCATATCATCTCCAGATCTCATCCCTTCCTCCGTTGGTGGCTTAC 1583

Gly 535 TyrIleGluAspAspGlyGlnIleIleSerAspValIeuLysIleProValGln----- 552

Db 1584 TACCAAGTGGAAATTAACGAATATTGTGCTGATTTCTGTGGGGAGATGGAAGATAC 1643

Gly 553 -----LeuValPheLys-----AsnLysIleLysIeuYr 562

Db 1644 TGCATGGGAACGTTGGTTGTAAGAGAGCGTCTCCAGAGACGATTCGAATCAAA----- 1697

Gly 563 TyrSerIysValLysAlaGluProSerGluLysValSerIeuAsnGlyIleSerValThrGln 582

Db 1698 -----AMCGCCAGAGCTGCATTAAGAAATCAAAATTCGAA--GGGGAT 1736

Gly 583 ProAspSerIleValaGlyIleValAlaValaLysIysSerValaAsnIleMetAsnAlaSer 602

Db 1737 CCAAGTGTGCGGTTGTGCTTGTGGCTGTGGACAAGAGATATATGTTCTCAATCATATAA 1796

Gly 603 AsnAspIleThrMetGluAsnValaHisGluLeuGlnIeuYrAsnThrGlyTyr 622

Db 1797 TATTAAGATTACCAAGCTAAGATATGGACACATATAGAAAGAGTGAATTTGGCTGTACA 1856

Gly 623 LeuGlyMetPheMetAsnSerPheAlaValaPheGlnGluCysGlyIeuTyrValIeuThr 642

Db 1857 GCTGCGAGTGGCCAGAAATTAATCTGGGTGTGTTGAAGATGCTGGATCGCTGTGCAACC 1916

Gly 643 AspAlaAsnIeuThr----- 647

Db 1917 AGCACTATATCTCAACACCAACACAGAGATCAGCTGCMAAGTGTCTCAGCTGCMAATCGG 1976

Gly 647 ----- 647

Db 1977 AGCGGTGACAGTTCGTGTTTCTGCTGTGACAGCAAGCAAGCAAAAGCGGACAGTTTCAG 2033

Gly 648 -----LysAspTyrIleAspGlyValYrAsnAsn----- 657

Db 2037 GATCAAGCGCTCGCTTAATGCTGTGAAGATATGTCATGCAAGAACCCCATGGGCTACCT 2096

Gly 658 AlaGluTyrAlaGluArgPheMetGluGluAsnGluGly----- 670

Db 2097 TGTGAAAAAGCTGCAAAATATATCATCAGAGAGGAATGCTTTAAGGCTGCTTCTTGA 2156

Gly 671 -----HisIleVal----- 673

Db 2157 TGTGTCTACATCAACAGGATCCGAGATGATAAACAACGAGAGAGGAGTTTCTG 2216  
 Qy 674 -----AspIleHisPheSerLeuGlySerSerProHisVal---ArgLysHis 689  
 Db 2217 GCAGAAAGTATTTTGAAGATGAAGTCTTTGGAGATGACAAACATCATCTCCAGGCTTGAT 2276  
 Qy 690 PheProGluThrTrpIleTrpLeu-----AspThrAsnMetGly 702  
 Db 2277 TTTCCTCGAGAGTGGTGTGGCTAAACAGAGAAATTGACCGGGAGACCTTAACATCAAGAGG 2336  
 Qy 703 TyrArgIleTyrGlnGluPheGluValThrValPro-----AspSerIleThr 718  
 Db 2337 ATTTCA-----AGCAAGACAGTACTTTTATCTGAGGAGATTCCATCA 2391  
 Qy 719 SerTrpValAlaThrGlyPheValIleSerGluAspLeuGlyLeuGlyLeuThrThr 738  
 Db 2382 ACCTGG-----GAGTTGCTGGCTGGCTGGCTTTCCACC 2417  
 Qy 739 -----ProValGluLeuGlnAlaPheGlnAlaPheGlnProPheIlePhe 751  
 Db 2418 AAGGAGATCTGTGGCTGAGACCGGTATGATAATACAGTATGAAAGACTTCTTCATTGAT 2477  
 Qy 752 LeuAsnLeuProTyrSerValIleArgGlyGluGluPheAlaLeuGluIleThrIlePhe 771  
 Db 2478 CTTGACGCTGACATTTTCAGTAGTAGAGATGAGAGAGTGGAGATTGAGCTATTCTGTAC 2537  
 Qy 772 AsnTyrIleuLysAspAlaThrGluValIleValIleIleGluLysSerAspLysPheAsp 791  
 Db 2538 AACCTGCTGACGAGATTTATGTGCGAGTGCAGTGCATGATATACACCCAGCTTC--- 2594  
 Qy 792 IleLeuMetThrSerSerGluIleAsnAlaThrGlyHisGlnGlnThrLeuLeuValPro 811  
 Db 2595 -----TCAGTGGCTTCCACAGAGACAGAACAGATACGACAGAGCTCCCAATTAA 2645  
 Qy 812 SerGluAspGlyAlaThrValLeuPheProIleArgProThrHisLeuGly-----Glu 829  
 Db 2646 GCCCTGCTCCAGACAGTACCATTTGTGTAGTCCCATTAAGACAGAGATTGCATGAT 2705  
 Qy 830 IleProIleThrValThrAlaLeuSerProThrAlaSerAspAlaValThrGluMetIle 849  
 Db 2706 GTTAGGTTATAGCAAGTGTCCGGGAGAGATTGGCATGATGCTGTGAGAGAAAGAACTG 2765  
 Qy 850 LeuValLysAlaGluGlyIleGluLysSerTyrSerGlnSerIleLeuLeuAspLeuThr 869  
 Db 2766 AAGGTTGACCTGAAGGGGAAACGAAATAATTGATCATTAATTATTAAGTACCAAGT 2825  
 Qy 870 AspAsnAlaGluGlnSerThr-----LeuLysThrIleuSerPheSerPheProAsn 887  
 Db 2826 GTAAAAGAGATTGTGGAACCCAGGAACTAAACGTCATAGCCCAATTAATTAGATGACAA 2885  
 Qy 888 ThrValThrGlySerGluArgValGlnIleThrAlaIleGlyAsp-----Val 903  
 Db 2886 GTGCTGTATACGAAGTTGAGAACCCAGGATTTCTGTGTAGGTGACCTGTGGCTCAGATT 2945  
 Qy 904 LeuGlyProSerIleAsnGly-----LeuAlaSerLeuIleArgMetProTyrGlyCys 921  
 Db 2946 ATTGAACATCATTTGATGAGAAATTAATCAATCATCATCATTAATTACTCTTGGCGCT 3005  
 Qy 922 GlyGluGlnAsnMetIleAsnPheAlaProAsnIle-----TyrIleLeuAspTyr 938  
 Db 3006 GGGGAGCAAAATATGATCATCATGATCATCGTCACTGACCACTTACTACTGAGACCA 3065  
 Qy 939 LeuThrLysLysGlnLeuThrAspAsnLeuLysGluValAlaLeuSerPheMetArg 958  
 Db 3066 ACAGGCGAGTGGGAATCTTGTGTGATGCGAGGACTGAAGCATATAAACAGATCTAG 3125  
 Qy 959 GlnGlyTyrGlnArgGluLeuLysTyrGlnArgGlySerPheSerAlaPheGly 978  
 Db 3126 ACTGTTATGCCAGACATGCTGTACAAAGAGAGATCTTCTATGACGACATTTACA 3185  
 Qy 979 AsnTyrAspProSerGlySerThrTrpLeuSerAlaPheValLeuArgCysPheLeuGlu 998  
 Db 3186 AAC-----CGTGATCTAGTTCTTGCTTAACAGCATATGTGTGAAGTCTTAGCCATG 3239

Qy 999 AlaAspProTyrIle-----AspIleAspGlnAsnValLeuHisArgThrTyrThrTrpLeu 1017  
 Db 3240 GCTTCCAAACATGATTAAGACATTTAGCAATGAATATTATTTGTGAGGTGAAATGGCTC 3299  
 Qy 1018 ---LysGlyHisGlnLysSerAsnGlyGluPheThrAspProLysArgValIleHisSer 1036  
 Db 3300 ATTCTGAACAGCAACACAGATGAGTGTCAAAAGAAATGCCCCGTGATCCATGGA 3359  
 Qy 1037 GluLeuGlnGlyValAsnLysSer-----ProValThrLeuThrAlaTyrIleVal 1053  
 Db 3360 GAATGCTGGAGAGAACTTAAGGCTGTAACCAAGACATCTTAAACAGCATTCATTGTG 3419  
 Qy 1054 ThrSerLeuLeuGlyTyrArg-----LysTyrGlnProAsnIle-----Asp 1067  
 Db 3420 ACTGCATTAATTGGAATCCAGATCATGCTGCAGAAAGAACAAATCATATTCTACAGACAGC 3479  
 Qy 1068 ValGlnGluSerIleHisPheLeuGluSerGluPheSerArgGlyIleSerAspAsnTyr 1087  
 Db 3480 ATCAATAGGCCACAGATTTATTTCAAAAGATGAGAAA---CTGCAAGGCTTTAC 3536  
 Qy 1088 ThrLeuAlaLeuIleThrTyrAlaLeuSerSerValGlySerProLysAlaLysGluAla 1107  
 Db 3537 ACTACAGCCCTCAAGCTATGCTTTGGCT-----GCTGCAGACCGA 3578  
 Qy 1108 LeuAsnMetLeuThrTrpArgAlaGluGlnGlyGlyMetGlnPheThrValSerSer 1127  
 Db 3579 CTCAATATGACAGGATCACTAGGACGATCAACAGAAAGAAATGGTTG----- 3629  
 Qy 1128 GluSerLysLeuSerAspSerTrpGlnProArgSerLeuAspIleGluValAlaAlaTyr 1147  
 Db 3630 -----GAAATATTAATGCTCGCACCATTAATTTGAGACGACCTTCAT 3674  
 Qy 1148 AlaLeuLeuSerHisPheLeuGlnPheGlnThrSerGluGlyIleProIleMetArgTrp 1167  
 Db 3675 GCTTGTGGCCCTGCTGAAATGAAGAAATTTGCTGAGGTGCGCTGTAGTCAGATGG 3734  
 Qy 1168 LeuSerArgLysArgAsnSerLeuGlyGlyPheAlaSerThrGlnAspThrThrValAla 1187  
 Db 3735 CTGATGATGCAAAATTTATGGGGAGACATATGACAAACCCAGCAACAGTTATGCTG 3794  
 Qy 1188 LeuValAlaLeuSerGluPheAlaAlaLeuMetAsnThrGluArg---ThrAsnIleGln 1206  
 Db 3795 TTTCAGGCTCTGCTGAATATGAGATTGATCAGATCCATCCATCAGACCTTAATTTGAT 3854  
 Qy 1207 ValThrValThrGlyProSerSerProSerPro-----LeuAla 1219  
 Db 3855 ATTTCTATTAACTGCCAGAACGAAAGTACCTGAAGGTAACGATTAATGATGAAT 3914  
 Qy 1220 ValValGlnProMetAlaVal-----AsnIleSerAlaAsn 1231  
 Db 3915 GCTGTCCAGGCCCGGACAGTATGAGACCAACTCAACGAACATCTCACTGTGTGATCA 3974  
 Qy 1232 GlyPheGlyPheAlaIleCysGlnLeuAsnValValTyrAsnValLysAlaSerGlySer 1251  
 Db 3975 GGTGATGAAAGCAACAAATGACCATTTGACGCTCTATATATACACATATG----- 4025  
 Qy 1252 SerTrpArgArgArgSerIleGlnAsnGlnGluAlaPheAspLeuAspValAlaVal--- 1270  
 Db 4026 ---AGGAGAGATCAAAATGTTTCACMAA---TTCCATCTTGATGTTTCTGTGAA 4076  
 Qy 1271 -----LysGluAsnLysAspAspLeuAsnHisValAlaAspLeuAsnVal 1284  
 Db 4077 AACGTGAATTTGAATTAACACGCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 4136  
 Qy 1285 CysThrSerPheSerGlyProGlyArgSerGlyMetAlaLeuMetGluValAlaLeuLeu 1304  
 Db 4137 TGACATGATGATCGGAGAAAGTTGATTTCAATATGACATTAATATGATATTTCTACTCTG 4196  
 Qy 1305 SerGlyPheMetValProSerGluAlaIle-----SerLeuSerGluThr 1319  
 Db 4197 ACTGTTTTCCTGATGCTGAAGACCTTAAAGGCTTTCTAACGAGTGAACAGATAC 4256

QY 1320 ValIysIyValGIuYrAspHis-----GlyIyLeuAsnLeuYrLeu 1334  
 Db 4257 AATCCCAAGTTTGAATTGCAATATATGCTCAGAAAGAACTGTGTCTATTTACTTA 4316  
 QY 1335 AspSerValAsnGIuYrGln---PheCysValAsnIleProAlaValArgAsnPhelys 1353  
 Db 4317 GACAGAGTCTCCCACTCTGAAGATGATATGCTGACATTTAAGATTACAAAGCATTTTGA 4376  
 QY 1354 ValSerAsnThrGlnAspAlaSerValSerIleValAspYrYrGIuProArgArgGln 1373  
 Db 4377 GTTGGCTTATTCAGCCAGAGATCAGTCAAGGTGTACAGCTACATCTAGATGACAA 4436  
 QY 1374 AlaValAsnSerYrAsnSerGIuValIyLeuSerSerCysAspIeuCysSerAspVal 1393  
 Db 4437 TGTACCAAGTTCTACCATCCAGATTAAGAAACAGAGCTTCTCAATAGATATGTCATGCT 4496  
 QY 1394 GlnGIyCysArgProCysGIuAspGIyAlaSerGIySerHisHisHis----- 1409  
 Db 4497 AACATTGGCCAGATGACAGAAACCTGTTCTTGCTCAACGACGAGAAAGATTGAT 4556  
 QY 1410 -----SerSerValIlePheIlePheCysPhelys 1419  
 Db 4557 CTTCAATTACGAATTCAAAAAGCCTGCGGCAAAATGTGATTTATCTTACAAACCAAG 4616  
 QY 1420 LeuLeu-----TyrPheWeGIuLeu 1426  
 Db 4617 CTGCTTCGAATGAGAAAGAAAGATGATATGATATCTATTTATGATGATGTT 4667

## RESULT 4

US-08-793-126-2  
 ; Sequence 2, Application US/08793126  
 ; Patent No. 5849297  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Harrison, Richard Alexander  
 ; APPLICANT: Faries, Charles Timothy  
 ; TITLE OF INVENTION: MODIFIED HUMAN C3 PROTEINS  
 ; NUMBER OF SEQUENCES: 2  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: HALE AND DORR LLP  
 ; STREET: 60 State Street  
 ; CITY: Boston  
 ; STATE: MA  
 ; COUNTRY: United States of America  
 ; ZIP: 02109  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: Patentin Release #1.0, Version #1.30  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/793,126  
 ; FILING DATE: 07-FEB-1997  
 ; CLASSIFICATION: 424  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Baker, Hollie L.  
 ; REGISTRATION NUMBER: 31,321  
 ; REFERENCE/DOCKET NUMBER: 102286,377  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (617) 526-6000  
 ; TELEFAX: (617) 526-5000  
 ; INFORMATION FOR SEQ ID NO: 2:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 5056 base pairs  
 ; TYPE: nucleic acid  
 ; STRANDEDNESS: single  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: cDNA  
 ; US-08-793-126-2

Alignment Scores:  
 Pred. No.: 4,44e-88 Length: 5056  
 Score: 931.00 Matches: 381  
 Percent Similarity: 40.88% Conservative: 287

Beet Local Similarity: 23.32% Mismatches: 609  
 Query Match: 12.67% Indels: 358  
 DB: 2 Gaps: 63

US-10-020-095-4 (1-1428) x US-08-793-126-2 (1-5056)

QY 3 GlyProProLeuLeuThrAlaIleHisLeuLeuCysValCysThrAlaIleVal 22  
 Db 65 GGTGCCACCTGCTG-----CTCTGCTACTTAACCCACCTCCCTGGCTCTG 112  
 QY 23 AlaProGlyProArgPheLeuValThrAlaProGlyIleIleArg----- 37  
 Db 113 GGG---AGTCCATGACTTACTATCAACCCCAACATCTTGCGGCTGAGAGCAGAG 169  
 QY 38 -----ProGlyIleValThrIleGlyValGIuLeuLeu 49  
 Db 170 ACCATGCTGTGAGGCCACGACCCGAGGGAGTTCAGTACTGTTCTGTCAC 229  
 QY 50 GIuHis-----CysProSerGIuValThrValIyAlaGIuLeuLeu 63  
 Db 220 GACTTCCCAAGCAAAAACTAGTGTCTGTCCAGTGAAGAGCTGTGACCCCTGCCAC 289  
 QY 64 LysThrAlaSerAsnLeuThrValSerVal---LeuGIuAlaGIuGIyValPheGIuYs 82  
 Db 290 AACCAATGGCGCAACGTCACTTCACGATCCAGCCCAAGAGGACTCAAGTCAAGAAAG 349  
 QY 83 GlySerPhelysThrLeuThrLeuProSerLeuProLeuAsnSerAlaAspGIuIleYr 102  
 Db 350 GGGCGCAACAGTTCTGTACCGTCGAGCC-----ACCTTGGAGCCCAAGTGTG 400  
 QY 103 GIuLeuArgValThrGIyArgThrGlnAspGIuIleLeuPheSerAsnSerThrArgLeu 122  
 Db 401 GAGAAAGTGTGCTGTGCTGACGCTGACAGCGGTACTC----- 439  
 QY 123 SerPheGIuThrLysArgIleSerValPheIleGlnThrAspIyAlaLeuYrIySPro 142  
 Db 440 -----TTCATCCAGACAGACAAAGACATCTACACCCCT 472  
 QY 143 LysGIuGIuValIyPheArgIleValThrLeuPheSerAspPhelYrProYrIyThr 162  
 Db 473 GGCTCCACAGTCTCTATCGATCTTCACTTCAACCCGCAACCAAGCTGCTACCGTGGCCGG 532  
 QY 163 SerLeuAsnIleLeuIleLysAspProLysSerAsnLeuIleGlnGln---TryLeuSer 181  
 Db 533 ACCGTCAATGTCATCAATTGAGAACCCGAGAGCATCCCGTCAACAGACTCTTGTCT 592  
 QY 182 GlnGlnSerAspLeuGIyValIleSerIyThrPheGlnLeuSerSerHisProIleLeu 201  
 Db 593 TCTCAGAACAGCTTGCGGTCTTGCTTGCGGACATTCGGAACCTGCTCAACATG 652  
 QY 202 GIyAspTTPSerIleGlnValGlnValAsnAsp-----GlnThrYrYrGIuSer 218  
 Db 653 GGCAGTGAAGATCCAGCTTACTATGAATCAACACAGCAGAGCTTCTCCACTGAG 712  
 QY 219 PheGlnValSerGIuYrValIleuProLysPheGIuValThrLeuGln-----ThrPro 236  
 Db 713 TTTAGGTGAGAGATGATGCTGCTGCCAGTTTCCAGGTCAATAGTGAAGCTTACAGAGAA 772  
 QY 237 LeuYrCysSerMetAsnSerIyHisLeuAsnGIyThrIleThrAlaLysYrThrYr 256  
 Db 773 TTCTACTACATCTATTAACGAGAGGGCTTGAGGTGCACATCACCAGCGAGTCTCTAC 832  
 QY 257 GIyIyProValIyGIyAspValThrIleuThrPhe----- 268  
 Db 833 GGAAGAAAGTGAAGGAACTGCTTGTCACTTCCGAGATCCAGATGCGCAACAGAG 892  
 QY 269 -----LeuProLeuSerPheTrGIyIyLysIyLysAsn 279  
 Db 893 ATTTCCCTGCTCGAATCCCTCAAGCGCATTCGATTTAGAGATGCTCGGGAGAGTTGTG 952  
 QY 280 IleThrLysThrPheLysIleAsnGIySerAlaAsnPheSerPheAsnAspGIuIuMet 299  
 Db 953 CTGACCCGAAAGTACTGCTGAGCGGGGTGCAAGACCCCGAGCAGAAAGACTGTGGGG 1012

QY 300 |yaaenValMeAapSerSeranelyLeuserGluTyrlLeuaPleuserSerProgly 319  
|:::|  
Db 1013 AAGTCTTTGTACGTCTGC----- 1033

QY 320 ProValGlulleThrThrValThrgluserValThnglyIleSerArgAnValSer 339  
|:::|  
Db 1034 -----ACCOTCACTTCGGCACTCAGGCAGTGATGATGTGCAGGCAAGCC 1078  
|:::|

QY 340 ThrAnVal PhePhelyleglnIhaApTyrlIeIlegluPheAspTyrlThrVal 359  
|:::|  
Db 1079 AGCGGATGCCCATCGTAGACTCTCCCTCACAGATCCACTTACCACAAGACCCAAATAC 1138  
|:::|

QY 360 LeulyPProSerLeuasnPhethrAlathrVallylsvalThrArgAlaApglyAsn-- 378  
|:::|  
Db 1139 TTCAAACCAAGAATACCCCTTGACCTCACTGCTTCGAGCAACCCGAATGGCTCTCCA 1198  
|:::|

QY 379 -----GlnLeuThrLeuglnLuglnArgAnAsnValIalleThrValThrglnArg 396  
|:::|  
Db 1199 GCCTAACCAAGTCCCCTGGCAGTCCAGGCGGAGACAGTGCAGCTCTTAACCCAGGGA 1258  
|:::|

QY 397 AsnyrThrgluTyrlTrpserGlySerAnserGlyAanglnLyMetElualavalGln 416  
|:::|  
Db 1259 GATGGCGTGGCCAAACTCAGCATCAACACACACCCCAACAGAG----- 1303

QY 417 LysIleaenyTrThrValProglInserGlyThrPhelysIlegluPheProIlLeuglu 436  
|:::|  
Db 1304 -----CCCTTGAGCATACAGCGTGGCCGACAGAAAMACAGAGCTCTCG 1345  
|:::|

QY 437 AapSerSerGluLeuglnLeuylalaTyrlPheLeuglySerLySerSerMechalval 456  
|:::|  
Db 1346 GAGGCAAGACG-----GCTACCAAGACCATGAGCGTCTG 1381  
|:::|

QY 457 HisSerLeuPheLySerProSerLyThrTyrlIeglnLeulyThr----- 472  
|:::|

Db 1382 CCTACAGACACCGTGGGCACTCCAAACATTACCTGCATCTCTCAGTCTCACTACAGAG 1441  
|:::|

QY 473 ----ArgaspGluasnlIelysValglySerProphegluLeuValIalseryIAsn 490  
|:::|  
Db 1442 CTCACAGCCCGGGGAGACCTCAACTCTCTCTCGCAATGACCGGCCCCACAGAG 1501  
|:::|

QY 491 LysAargLeulysgluLeuserTyrmetyValIalsaraGlylglnLeuValIalvalgly 510  
|:::|  
Db 1502 GCCAAGATCCCGTACTACACCTCACTCATMGAAACAGGCGAGCTTTGAAGCGGGA 1561  
|:::|

QY 511 Lys-----GlnAsnSerThmetPheSerLeuThrProgluasner 524  
|:::|  
Db 1562 CGCCAGGTCGAGAGCCCGCGCAGAGACTCGTGtGtGtCCCCCTTCATCAACACGAC 1621  
|:::|

QY 525 TrpThrProLySAlaCyValIleValTyrlTyrl-----IlegluAspApgly 540  
|:::|  
Db 1622 TTCATCCCTTCCCTCGCGCTGGCGGCGATCAACGCTGATCCGAGCCAGGCGCAGAG 1681  
|:::|

QY 541 GlulleIleSerAspValleuLyIleProValGln-----LeuVal 554  
|:::|  
Db 1682 GAGGTGTGGCCGACTCCGTGGGTGAGCGTCAAGAGACTCCGTGCGTGGCTG 1741  
|:::|

QY 555 PheLyAsnLySilleLyLeuTyrlTrpserLySValLySAlaglu----- 569  
|:::|  
Db 1742 GTAAAAGCGGC-----CAGTCAAGAGACCGGACGCTGTA 1777  
|:::|

QY 570 ProserGluLyValSerLeuArgIleSerValThrglnProaspSerIleValglyIle 589  
|:::|  
Db 1778 CTTGGGACGACATACCTCTGAAGTATAGAGGTACACAGGGGCCCGGGGTGTA---CTG 1834  
|:::|

QY 590 ValAlaValAspLySSerValAsnLeuMetAnlaSerAsnAspIleThrMetcluAsn 609  
|:::|  
Db 1835 GTGGCCGTGACAAAGCGCGTGTCTGTGTAATAAGAAACAAACTGACCAAGATNAG 1894  
|:::|

QY 610 ValValHISgluLeugluLeuTyraenthrglyTyrlLeuglyMetPheMechanser 629  
|:::|  
Db 1895 ATCTGGACGTGTGGAGAGGACAACTCGCTTCACCCCGGACAGTGGAAAGATTAC 1954  
|:::|

QY	630	PheAlaValPheGlnGluCysGlyLeuTrpValLeuThrAspAla-----	644
DB	1955	GCCGGTGTCTTCCGACGCGAGGCTGACCTTCACGACGACGACAGTGGCCAGACGCGC	2014
QY	644	-----	644
DB	2015	CAGAGGCGAAGACTTCACTGCGCCGACGACGCCGCCGACGCCCTTCCTGCAGCTC	2074
QY	645	-----AsnLeuThrLysAspTrpLe	651
DB	2075	ACGGAGAAGCGAATGACAAAGTCGCGCAATGTAACCCAAAGAGACTGCGCAATGCTGCGAG	2134
QY	652	AspGlyValTrpAspAsn-----AlaGluTrpAlaGluArgPheMetCys	666
DB	2135	GACGCGATGCGCGAGAAACCCCATGAGTTCCTGTCGACGCGCCGGAACCCGTTTCACTCC	2194
QY	667	GluAsnGlu-----GlyHisIleValAspIleHis	676
DB	2195	CTGGGCGAGGCGTGCAGAAGGTTCTCTGGACTGCTGCACATCAATCAAGAGCTGGGG	2255
QY	677	-----AspPheSerLeu	680
DB	2255	CGCGACGACGCGGGGCGCAGCCACTGCGCTGGCCAGAGTAACCTGGAATGAGAGCACTC	2314
QY	681	GlySerSerProHisVal-----ArgLysHisPheProGluThrTrpIleTrp	696
DB	2315	ATTGCAAGAGAACAATCGTTTCCCGAAGTAGTTCACAGAGAACTGCGCTGGAACGTT	2374
QY	697	-----LeuAspThrAsnMetGlyTrpArgIleTrp	706
DB	2375	GAGCACTTGAAAGGCCACCGCAAAATGGAATCTCAAGAACCTC-----	2419
QY	707	GlnGluPheGluValThrValProAspSerIleThrSerTrpValAlaThrGlyPheVal	726
DB	2420	-----ATGAATATAATTTTGAAGAACTCCATCAACAGTGGAGAAATCTGCGTGAAC	2479
QY	727	IleSerGluAspLeuGlyLeuGlyLeuThrThrTrpProValGluLeuGlnAlaPheGln	746
DB	2474	ATGTCGACAAAGAAAGGATC-----TGTGTGGCAGACCCCTTCGAGCGACAGTAATGACG	2530
QY	747	ProPhePheIlePheLeuAsnLeuProTyrSerPaiIleArgGlyGluGluPheAlaLeu	766
DB	2531	GACTTCTTCAATCAGCTCGCGCTACCTCACTCTGTGTGCAGAAAGAGACAGTGGAAATC	2590
QY	767	GluIleThrIlePheAsnTrpLeuLysAspAlaThrGluValLysValIleIleGlu	785
DB	2591	CGACCGCTTCTCAATTAC-----CGGCAGAACAAGAGCTCAAGGTGAGGTGAACCTA	2644
QY	786	---LysSerAspLysPheAspIleLeuMetThrSerSerGluIleAsnAlaThrGlyHis	804
DB	2648	CTCCACAATCCAGCTTCTGAGCCTGAGCCACCAACAAAGAG-----CGTCAC	2695
QY	805	GlnGlnThrIleLeuValProSerGluAspGlyAlaThrValLeuPheProIleArgPro	824
DB	2696	CAGCAGACAAATMACATCCGCCCAAGTCTCGTTGCTCCTTCATATGTCATGTGCGG	2755
QY	825	-----ThrIleLeuGlyGluIleProIleThrValThrAlaLeuSerProThrAlaSer	842
DB	2756	CTAAGAAGCCGCGCTGCAGAGAGTGAAGTCAAGCGCTGCTCAATCATTTACATCACT	2815
QY	843	AspAlaValThrGlnMetIleLeuValLysAlaGluGlyIleGluTyrSerTrpSerGln	862
DB	2816	GACGGTGCAGGAAGTCCCTGAAGTGTGTGCGGAAGAAATCAAGATGAACAAAACTGTG	2875
QY	863	SerIleLeuLeuAspLeuThrAspAsnArgLeu-----GlnSerThrLeuLysThrLeu	880
DB	2876	GCTGTT-----CGACCCCTGATCCAGAAACGCTGCGCGCTGAGAGAGTCAGAAAAG--	2925
QY	881	SerPheSerPheProProAsnThrVal-----ThrGlySerGluArg	894
DB	2930	-----GACATCCCACTCGACAGCTCAATGACCAAGTCCCGGACACCGAGATCTGAGACC	2983
QY	895	-----ValGlnIleThrAlaIleGlyAspValLeuGlyProSerIleAsnGly	910



```

REFERENCE/DOCKET NUMBER: 102286.377
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 526-6000
TELEFAX: (617) 526-5000
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 5056 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-09-132-271-2

Alignment Scores:
Pred. No.: 4,44e-88 Length: 5056
Score: 931.00 Matches: 381
Percent Similarity: 40.88% Conservative: 287
Best Local Similarity: 23.32% Mismatches: 609
Query Match: 12.67% Indels: 358
DB: 3 Gaps: 63

US-10-020-095-4 (1-1428) x US-09-132-271-2 (1-5056)
QY 3 GlyProPheLeuThrAlaHisLeuLeuCyValCysThrAlaAlaLeuAlaVal 22
   |||||
DB 65 GGTCCAGCCTGCTG-----CTCCTGCTACTAACCCACCTCCCTGCTGCTG 112
QY 23 AlaProGlyProAlaGlyPheLeuValThrAlaProGlyIleIleArg----- 37
   |||||
DB 113 GGG---AGTCCCATGACTCTATCATCAACCCCAACCTTGTGCGCTGAGAGCGAGAG 169
   |||||
QY 38 -----ProGlyGlyAsnValThrIleGlyValGluLeuLeu 49
   |||||
DB 170 AACATGATGCTGAGAGCCGACGAGCGGAGGATGTTCCAGTCACTGTTACTGTCCAC 229
   |||||
QY 50 GluHis-----CysProSerGlnValThrValIleValIleGluLeuLeu 63
   |||||
DB 230 GACTTCCCGAGCAAAAACTAGTGTCTGTCCAGTGAGAGACTGTGCTGACCCCTGCCACC 289
   |||||
QY 64 LysThrIleAsnLeuThrValIleSerVal---LeuGluAlaGluGlyValPheGluLys 82
   |||||
DB 290 AACCATGCGGCAACGTCATCTTCAACATCCACGACCAACAGGAGTTCACTGACGAAAG 349
   |||||
QY 83 GlySerPheLysThrLeuThrLeuProSerLeuProLeuAsnSerAlaAspGluIleTyr 102
   |||||
DB 350 GGGGCAACAAGTTCGTGACCGTGCAGGCC-----ACCTCGGAGCAACGAGTGTG 400
   |||||
QY 103 GluLeuArgValThrGlyArgThrGlnAspGluIleLeuPheSerAsnSerThrArgLeu 122
   |||||
DB 401 GAGAAAGTGTGTGCTGCTGACGCTGACGAGCGGCTACCTC----- 439
   |||||
QY 123 SerPheGluThrLysArgIleSerValPheIleGlnThrAspLysAlaLeuTyrLysPro 142
   |||||
DB 440 -----TTCAATCCAGACAGACAGACCAATCTACACCTT 472
   |||||
QY 143 LysGlnGluValLysPheArgIleValThrLeuPheSerAspPheLysProTyrLysThr 162
   |||||
DB 473 GGGTCCACAGTTCCTATCGGATCTTCCACCTCAACCAACAGCTGACCTGCGGGCGG 532
   |||||
QY 163 SerLeuAsnIleLeuIleLysAspProLysSerAsnLeuIleGlnGln---TyrLeuSer 181
   |||||
DB 533 ACGGTCATGTCATCAATGAGAACCCGGAGGACATCCCGGTCAAGACGACTCTTGTCT 592
   |||||
QY 182 GlnGlnSerAspLeuGlyValIleSerLysThrPheGlnLeuSerSerHisProIleLeu 201
   |||||
DB 593 TCTAGAACCAAGCTTGGGCTTGTCCCTTGTCTTGGGACATTCGGAACCTGTCAACATG 652
   |||||
QY 202 GlyAspTyrSerIleGlnValIleValAsnAsp-----GlnThrTyrTyrGlnSer 218
   |||||
DB 653 GGCAGAGTGAAGATCCAGCTTACTATGAAAATCAACACAGCAGGCTTCTTCCACTGAG 712
   |||||
QY 219 PheGlnValSerGluTyrValLeuProLysPheGlnValThrLeuGln-----ThrPro 236
   |||||

```

```

DB 713 TTGAGGTGAAGAGTACGTGCTGCCAGTTTCGAGCTCATGTGAGCTTACAGAGAA 772
QY 237 LeuTyrCysSerMetAsnSerLysHisLeuAsnGlyThrIleThrAlaLysTyrThrTyr 256
   |||||
DB 773 TTCTACTACATCTATTAACGAGAGGGCTGTGAGTGCATCATCCCGCAGGTTCTCTAC 832
   |||||
QY 257 GlyLysProValIleGlyAspValThrLeuThrPhe----- 268
   |||||
DB 833 GGGAGAAAGTGAAGGAACTGCTTGTTCATCTTGGGATCCAGATGCGGAACAGAG 892
   |||||
QY 269 -----LeuProLeuSerPheThrGlyLysLysAsn 279
   |||||
DB 893 ATTTCCCTGCTGAATCCTCAAGCCGATTCGATGAGAGATGCTCGGGGAGTTGTG 952
   |||||
QY 280 IleThrLysThrPheLysIleAsnGlySerAlaAsnPheSerPheAsnAspGluMet 299
   |||||
DB 953 CTGAGCCGGAAGTACTGCTGAGCGGGGTGCACAAACCCCGACAGAAAGCCTGTGGGG 1012
   |||||
QY 300 LysAsnValMetAspSerSerAsnGlyLeuSerGluTyrLeuAspLeuSerProGly 319
   |||||
DB 1013 AAGTCTTTGACGTGCTGCC----- 1033
   |||||
QY 320 ProValGluIleLeuThrThrValThrGluSerValThrGlyLysSerArgAsnValSer 339
   |||||
DB 1034 -----ACGTCATCTTGACCTCAGGACGTACATGTGTCAGCGACAGCCG 1078
   |||||
QY 340 ThrAsnValPhePheLysGlnHisAspTyrIleIleGluPhePheAspTyrThrThrVal 359
   |||||
DB 1079 AGCGGATCCCACTGCTGACTCTCCCTTACCAATCCTTCCACCAAGACCCCAAGTAC 1138
   |||||
QY 360 LeuLysProSerLeuAsnPheThrAlaThrValLysValThrArgAlaAspGlyAsn--- 378
   |||||
DB 1139 TTCAAACAGAGATGCCCTTGTGACTCATGTGTGTGTCAGCAACCTGATGCTCTCCA 1198
   |||||
QY 379 -----GlnLeuThrLeuGluIleArgArgAsnAsnValIleThrValThrGlnArg 396
   |||||
DB 1199 GCCTACCAAGTCCCGGTGAGTCAGCGGAGGAGACCTGTGCATCTTCAACCGAGGA 1258
   |||||
QY 397 AsnTyrThrGluTyrTrpSerGlySerAsnSerGlyAsnGlnLysMetGluAlaValGln 416
   |||||
DB 1259 GATGGGTGGCCAACTCAGATCAACACACACCCCAAGCCAGAG--- 1303
   |||||
QY 417 LysIleAsnTyrThrValProGlnSerGlyThrPheLysIleGluPheProIleLeuGln 436
   |||||
DB 1304 -----CCCTGACATCAACGCTGCGCAGCAAGAACGAGAGCTCTCG 1345
   |||||
QY 437 AspSerSerGluLeuGlnLeuLysAlaTyrPheLeuGlySerLysSerMetAlaVal 456
   |||||
DB 1346 GAGGACAGACAG-----GTTACCGAGACCATGACAGGCTCTTG 1381
   |||||
QY 457 HisSerLeuPheLysSerProSerLysThrTyrIleGlnLeuLysThr----- 472
   |||||
DB 1382 CCTTACAGCACCGTGGGCAACTCAACATTTACTGTGATCTTCACTGCTTACGTACAGAG 1441
   |||||
QY 473 -----ArgAspGluAsnIleLysValIleGlySerProPheGluLeuValIleSerGlyAsn 490
   |||||
DB 1442 CTCAGACCGGGGAGACCTCAACGTCATCTTCTCTGGAATGAGACCGGCCACAGAG 1501
   |||||
QY 491 LysArgLeuLysGluLeuSerTyrMetValIleSerArgGlyGlnLeuValAlaValGly 510
   |||||
DB 1502 GCCAAATCCGCTACTACACTTACCTGATCATGAAAGGCGAGCTGTGGAAGCGGGA 1561
   |||||
QY 511 Lys-----GlnAsnSerThrMetPheSerLeuThrProGluAsnSer 524
   |||||
DB 1562 CGCCAGGTGAGAGCCCGGCGCAGAGCTGTGTGTCTGCTGCTGCTGCTGCTGCTGCTG 1621
   |||||
QY 525 TrpThrProLysAlaCysValIleValIleValTyrTyr-----IleGluAspArgGly 540
   |||||
DB 1622 TTCAATCCTTCTTCCGCGCTGTGAGGTACTACAGCTGATGTGTGTCAGCGCCAGAG 1681
   |||||
QY 541 GluIleIleSerAspValLeuLysIleProValGln-----LeuVal 554
   |||||
DB 1682 GAGGTGTGTCGACATCCGTGTGTGTGAGCTCAAGAGACTCTGCTGTGAGCTGTGCTGTG 1741
   |||||

```





```

Db      3752  TATGCCCTCTTGCC---CTATGAGAGCTAAAGACTTTGACTTTGTGCTCCCGCTG 3808
Qy      1166  ArgTPLLeuSerArgGlnArgAhnSerLeuGlyGlyPheAlaSerThrGlnAspThr 1185
Db      3809  CGTTGGCTCAATGAACAAGAGATACTACAGCGTGGTGGCTATGCTTACACAGGCCACTTC 3868
Qy      1186  ValAlaLeuLeuAlaLeuSerGluIuheaIala---LeuMetAsnThrGluArgThrAsn 1204
Db      3869  ATGGTGTTCCAAAGCCCTTGCTCAATAATCAAAAAGAGCGCCCTGACACCAAGAACTGAAC 3928
Qy      1205  IlegInValThrValThrGlyProSerSer-----ProSerProleu 1218
Db      3929  CTGTATGTGTCTCCCTCACTCACTGCCACGCG-CAGCTCCAAAGATCAACCCAGTATCCACTG 3987
Qy      1219  AlaValValGlnPrometAlaVal-----Asn-IleSe 1229
Db      3988  GGAATCTGCCACGCTCTGCGCATCAGAAGAGACCAGAAATGAGGGTTTCACAGTAC 4047
Qy      1229  rAlaasnGlyPheGlyPheAlaIleCysGlnIleuAsnValValTyraAsnValValAlaSe 1249
Db      4048  AGCTAAAGAAAAAGGCCAAAGCACCTTGCGTGGTGTCAGTAAGTACCAATGACCTAAGGCCAA 4107
Qy      1249  rGlySerSerArgArgArgSerIleGlnAsnGlnGluAlaPheAspLeuAspValAl 1269
Db      4108  AGATCAACTCACTCACTGTAATAAA-----TTGACCTCAAGTAC 4166
Qy      1269  aVal-----LysGlnuAsnLysAspAspLeuAsnHisValAspLe 1282
Db      4147  CATAAACCAAGCACCGGAACAGAAAGAGGCGCTCAGATGCCAAGAACATATGATCCT 4206
Qy      1282  uAsnValCysThrSerPheSerGlyProGlyArgSerGlyMetAlaLeuMetGluValAs 1302
Db      4207  TGAGATCTGTATCCAGATACCGGGAGAACCAAGATGCCATATGTCTTAATTTGACATATC 4266
Qy      1302  nLeuLeuSerGlyPheMetValProSerGlu-----AlaIleSerLeuSe 1317
Db      4267  CATGTAGTACTGGCTTTGCTCCACAGACACAGATGACTGAAAGACGCTGGCAATGGTGTG 4326
Qy      1317  rGluThrValLysLysValGluTyT-----AspHisGlyLysLeuAsnLe 1332
Db      4327  CAGATACATCTCCAGATATGAGCTGAGCAAAAGCTTCTCCAGATGGAAACACCTCATCAT 4386
Qy      1332  uTyTLeuAspSerValAsnGluThrGln---PheCysValAsnIleProAlaValArgAs 1351
Db      4387  CTACCTGGAACAAGGCTCTCACTCACTGAGAGATGACTGTAGCTTCAAAAGTTCACCAATA 4446
Qy      1351  nPheTyValSerAsnThrGlnAspAlaSerValSerIleValAspTyTyTrGluProAr 1371
Db      4447  CTTTAATGATGAGCTTATTCACACTGAGCACTCAAGCTTACGCTTATTCACACTGGA 4506
Qy      1371  gArgGlnAlaValArgSerTyTAsnSerGluVal-----LysLeuSerSerCysAs 1388
Db      4507  GGAAGAGCTGACCGGTTCTTACATCCGGAAGAGAGATGGAAGAGCTGAACAAG----- 4561
Qy      1388  pLeuCysSerAspValGlnGlyCysArgProCysGluAsp 1401
Db      4562  -CTTGCCGTGAT---GAACTGTGCCGCTGTGCTGAGAG 4597

```

```

; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 23
; LENGTH: 5067
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-142-334-23

Alignment Scores:
Pred. NO.:      4,46e-88
Score:          931.00
Percent Similarity: 40.88%
Best Local Similarity: 23.32%
Query Match:    12.67%
DB:              3

Length: 5067
Matches: 381
Conservative: 287
Mismatches: 609
Indels: 358
Gaps: 63
US-10-020-095-4 (1-1428) x US-09-142-334-23 (1-5067)

```

QY 3 GlyProPheIeuLeuThrAlaAlaHisPheIeuCysValCySthrAlaAlaIleuValAla 22  
||||| |||||  
Db 76 GGrTCAGGCGCTGTG-----CTCCGTACTTAACCAACCCTCCCCTGGCTCTG 123

QY 23 AlaProGlyProArgPheLeuValThrlAlaProGlyIleIeArg----- 37  
||||| |||||  
Db 124 GGG---AGTCCATGTACTTACTTACTCATCACCCCCCAACATTGCGGGCTGGAGAGGAGGAG 180

QY 38 -----ProGlyIleAnValThrlIeGlyIleValIGluLeu 49  
||||| |||||  
Db 181 ACCATGGTGCTGGAGGCCAACAACCGGAGGGAGTGTTCAGTCACTGTTACTGTCCAC 240

QY 50 Gluhia-----CysProSerGlnValThrValLybAlaGluLeu 63  
::: |||||  
Db 241 GACTGCCAAGCAAAACATAAGTGTGTGCATGAAGAAGATGTGTGCAGCCCTGCAC 300

QY 64 LyethrAlaSerAsnLeuThrlAlaSerVal---IeuGlualagIuGlyValPheGluLys 82  
||||| |||||  
Db 301 AACCATGGGCAAGTCCTACCTTCACGATCCAGCAACAGGAGTTCAAGTCAGAAAG 360

QY 83 GlySerPheLysThrLeuThrlleuProSerIeuProIeuAsnSerAlaAspGluIeTy 1020  
::: |||||  
Db 361 GGAGGCAACAAGTTCGTGACCGTGCAAGGC-----ACCTTCGGAGAACCAAGTG 4111

QY 103 GluleuArgValThrclyArgThrglInaspGluIleuPheSerAsnSerThrArgLeu 1222  
||||| |||||  
Db 412 GAAGAAGTGCTGTGTGTGCAGCTGCAGAGCGGATCCTC----- 4500

QY 123 SerPheGluThrLysArgIleSerValPheIleGlnThraSpLybAlaLeuTyLysPro 1422  
||||| |||||  
Db 451 -----TTCAATCCAGACAGCAAGACATCAACCCCT 4833

QY 143 LyeGlnGluValLysPheArgIleValThrlleuPheSerAspPheLybProTyLysThr 1622  
||||| |||||  
Db 484 GGCTCCACAAGTTCATCGATCTTCACCGTCAACCAACAGCTACCCGTGGGCGCG 5433

QY 163 SerIeuAsnIleLeuLysPheProLysSerZanLeuIleGlnGln---TrpleuSer 1818  
::: |||||  
Db 544 ACAGTCATGTCAACTTAGAACCCGGAAGGCAATCCCGGTCAAGCAGACTCTTGTCT 6033

QY 182 GlnGlnSerAspLeuGlyValIleSerLysThrPheGlnIeuSerSerHispollIeu 2010  
::: |||||  
Db 604 TCTCAACAACAGCTTGCGCTGTGCGCTGTGTGGACATTCGGGAACCTGTCAACATG 6633

QY 202 GlyAspTrpSerIleGlnValGlnValAsnAsp-----GlnThryTyGlnSer 218  
||||| |||||  
Db 664 GGCACAGTGAAGATCCGACCTACTATGAAAATCAACACAGCAGAGTCTTCCACTGAG 7233

QY 219 PheGlnValSerGluTyValIleuProLysPheGluValThrlleuGln-----ThrPro 2366  
::: |||||  
Db 724 TTGAGAGTGAAGAGTACGTGCTGCCAGATTTCGAGGTCAATAGTGGAGCCACAGAGAAA 7833

QY 237 LeuTyCySerMetCAsnSerLysHisIleuLenglyThrlIethrAlaLysTyThrTy 2566  
::: |||||  
Db 784 TTCTACTTCAATCTTATACAGAAAGGGCTTGAGGTCAACATCAACCGGAGTCTCTAC 8433

OY	257	GlyLysProValIySGIAspValThrLeuThrPhe-----	268
Db	844	GCGAAGAAAGTGAAGGAACTCCCTTTGTATCTTCGGGATCCAGAGTGCACAACAGAG	903
OY	269	-----LeuProLeuSerPheTrpGIyLysIySylsaen	279
Db	904	ATTTCCTCCGCGTAATCCCTCAAGGCAATCCGATTGAGAGATGGCTCGGGGAGCTGTGG	963
OY	280	IleThrLyThrPheLysIleasnGLysAlaIasPheSerPheasnApGIuIuMet	299
Db	964	CTGACC CGGAAGTACTGCTGACGCGGGTGCAGAACCCCGCAGCAGAAGACTGTGGGG	1022
OY	300	LysAsnValMetAspSerSerasnGLyLeuSerGIyTyLeuAspLeuSererProGIy	319
Db	1024	AAGTTTTGTATACGTCTGCC-----	1042
OY	320	ProValGIuIleLeuThrThrValThrIuSerValThrGIyIleSerArgasnValser	339
Db	1045	-----ACCGTATCTTGCACTCAAGGAGTGAAGCATGTGTGACGACAGCGCC	1082
OY	340	ThrAsnValPhePheLysGLInHisaspTyrlleIleGIuIlePhePheaspTyrlThrVal	359
Db	1090	AGCGGATCCCATGTCGACTCTCCCTACAGATCACTTACCAACCAAGACCCAAAGTAC	1142
OY	360	LeuLysProSerLeuAsnPheThrAlatThrValIysValThrArgAlaaspIyAsn---	378
Db	1150	TTCAAACCAAGATACCCCTTTAACCCTCATGTGTTCGTACCAACCCGTATGCTCTCCA	1202
OY	379	-----GIuLeuThrLeuGIuGIuArgArgAsnValIleThrValThrGIuArg	396
Db	1210	GCCTAACGATCCCCCTGGCACTCCAGGCGCAGAGCACATGTGCATCTTAAACCCAGGGA	1262
OY	397	AsnTyrlThrGIyTyrlTrpSerGIySerAsnSerGIyAsnGIuLysMetGIuAlavalGIN	416
Db	1270	GATGCGGTGGCCAAACTCAATCAACACACACCCAGCCAGAG-----	1312
OY	417	LysIleAsnTyrlThrValProGIuIleGIyThrPheLysIleGIuIlePheProIleLeuGIu	436
Db	1315	-----CCCTTGAGCATCACGGTGGCGCAGCAAGAGCAGAGACTCTCG	1352
OY	437	AspSerSerGIuLeuGIuLeuLysAlaTyrlPheLeuGIySerIySerSerMetAlaVal	456
Db	1357	GAGGCAGAGCAG-----GCTAACGAGGCATGCAAGGCTCTG	1392
OY	457	HisSerLeuPheLysSerProSerLySThyrlleGIuLeuLysThr-----	472
Db	1393	CCCTACAGACCGCTGGGCACTCCAACATTAATCTGCACTCTCACTGCTACGTACAGAG	1452
OY	473	-----ArgAspGIuAsnIleLysValGIySerProPheGIuLeuValIasergIyaen	490
Db	1453	CTCAGACCCGCGGAGAACCTCAACGCACTTCTCTCCGAATGAGACCGGCCCCACAG	1512
OY	491	LysAsnGLyLeuGIuLeuSerTyrlMetValIasArGIyGIuLeuValIalvalGIy	510
Db	1513	GCCAAGATCCGCTACTACACCTCACTGATCATGAAMAAAGGAGCGCTGTGGAAGCGGGA	1572
OY	511	Lys-----GIuAsnSerThrMetPheSerLeuThrProGIuIuAsner	524
Db	1573	CGCCAGTGCAGAGACCCGCGCAGAGACCTGTGGTGTCTGCCCTGTCCATCAACACCAAG	1632
OY	525	TrylThrProLYsAlCyAsValIleValTyrlTyrl-----IleGIuAspApGIy	540
Db	1633	TTCATCCCTTCTCTCCGCTGGTGGGTACTACACGCTGATCGGTGCCAGCGGCAGAG	1692
OY	541	GIuIleIleSerAspValLeuLysIleProValGIN-----LeuVal	554
Db	1693	GAGGTGTGGCCGCACTCCGTGTGGGTGACGTCAAGAGACTCTGCGTGGGCTTCGTGGT	1752
OY	555	PheLysAsnValIleLysLeuTyrlTrpSerLySvalIyAlaGIu-----	569
Db	1753	GTAATAAGCGGC-----CACTCAGAAACCGCAGCCTGTGA	1782
OY	570	ProSerGIuLysValSerLeuArgIleSerValThrGINProAspSerIleValGIyIle	589

Db 1789 CCTGGGCGACGATGACCCCTGAAAGATAGAGGTACCACGGGGCCCGGGTGTGA---CTG 1845  
 Oy ValAlaValAspIysSerValAsnIleuMetAsnIleAsenAspIleThrMetGluAsn 609  
 Db 1846 GTGGCCCGGTGACAAAGGGGGGTGTTCGTGTGTAAATAAGAACAAACTACGACGAGTAA 1905  
 Oy 610 ValValHisGluLeuGluLeuIleuThrGlyThrGlyIleuGlyMetPheMetAsnSer 629  
 Db 1906 ATCTGGGACGTGTGTGAGAAAGGCGACAGATCGGTGCACCCCGGGCGAGTGGAAAGATTAC 1965  
 Oy 630 PheAlaValPheGlnGluCysGlyLeuTrpValLeuThrAspAla----- 644  
 Db 1966 GCCGGTGTCTTCCGACGCGAGGGCTGACTTCAAGACGACAGTGGCCGACGACCGCC 2025  
 Oy 644 ----- 644  
 Db 2026 CAGAGGGGAGAACTTCAGTGTCCCGACGACCGCCCGCGACCGCCGTCCTCGTGCAGCTC 2085  
 Oy 645 -----AsnIleuThrIysAspTrpIle 651  
 Db 2086 ACGGAGAAAGCGAATAGACAAAGTGTGCGAAGTACCCCAAGAGAGTCGCGAAATGTCTGCGAG 2145  
 Oy 652 AspGlyValIleYrAspAsn-----AlaGluTrpAlaGluArgPheMetGlu 666  
 Db 2146 GACGCGACATGCGGAGAAACCCCATGAGTTCTGTGCGACGCGCGGACCGGTTTCAATCC 2205  
 Oy 667 GluAsnGlu-----GlyHisIleValAspIleHis 676  
 Db 2206 CTGGGCGAGGGGTGCAAGAAGGTCTTCTGTGACTGTCTGCAACTACATCAGAGAGCTGGG 2265  
 Oy 677 -----AspPheSerLeu 680  
 Db 2266 CGGACGACACGCGCGGGCCAGCACCTGTGGCGTGGACGAGATAACTGGAATGAGAGACATC 2325  
 Oy 681 GlySerSerProHisVal-----ArgLysHisPheProGluThrTrpIleTrp 696  
 Db 2326 ATTCACGAAAGAACATCGTTTCCGAAAGTAGTTCACGAGACATGCGTGTGAAACGTT 2385  
 Oy 697 -----LeuAspThrAsnMetGlyTrpArgIleTrp 706  
 Db 2386 GAGGACTTGAAGAAGCCACCGAAAAATGGAATCTCTACGAAGCTC----- 2433  
 Oy 707 GlnGluPheGluValThrValProAspSerIleThrSerTrpValAlaThrGlyPheVal 726  
 Db 2431 -----ATGAATATATTTTGGAAAGACTCCATCAACAGTGGAGATTCTGGCTGTGAC 2484  
 Oy 727 IleSerGluAspLeuGlyLeuGlyLeuThrThrProValGluLeuGlnAlaPheGln 746  
 Db 2485 ATGTGCGACAAAGAAAGGGATC---TGTGTGCGACACCCCTTCGAGGTCAACAGTATGACG 2541  
 Oy 747 ProPhePheIlePheLeuAsnLeuProIleSerValIleArgGlyGluGluPheAlaLeu 766  
 Db 2542 GACTTCTTCATCGACCTGTGCGCTGACCTTACTCTGTGTGCAAAAGACAGAGTGGAAATC 2601  
 Oy 767 GluIleThrIlePheAsnTrpLeuLysAspAlaThrGluValLysValIleIleGlu 785  
 Db 2602 CGAACCCGTCTCTACATTAC---CGGAGAACCAAGAGCTCAAGGTGAGGTGGAACTA 2655  
 Oy 786 ---LysSerAspLysPheAspIleLeuMetThrSerSerGluIleAsnAlaThrGlyHis 804  
 Db 2659 CTCACCAATCATCAGCTTCTGTGACGCTGGCCACCAAGAG-----CGTAC 2706  
 Oy 805 GlnGlnThrLeuLeuValProSerGluAspGlyAlaThrValLeuPheProIleArgPro 824  
 Db 2707 CAGGAGACCAATACCATCTCCCCCAAGGCTCGTGTGTCGTTCCATATGTCATGTGCTG 2766  
 Oy 825 -----ThrIleGluGlyGluIleProIleThrValThrAlaLeuSerProThrAlaSer 842  
 Db 2767 CTAAAGACCGGCGCTGACGAGAAAGTGAAGTCAAGGCTGTGTCTACATCATTTATCATAGT 2826  
 Oy 843 AspAlaValThrGlnMetIleLeuValLysAlaGluGlyIleGluLysSerTrpSerGln 862



```

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/662,227
FILING DATE: 14-JUN-1996
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: OBLON, NORMAN F.
REGISTRATION NUMBER: 24,618
REFERENCE/DOCKET NUMBER: 1126-0107-0X
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-413-3000
TELEFAX: 703-413-2220
INFORMATION FOR SEQ. ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 5948 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-08-662-227-1

Alignment Scores:
Pred. No.: 5.24e-84 Length: 5948
Score: 894.00 Matches: 361
Percent Similarity: 40.90% Conservative: 295
Best Local Similarity: 22.51% Mismatches: 672
Query Match: 12.17% Indels: 276
DB: 2 Gaps: 56

US-10-020-095-4 (1-1428) x US-08-662-227-1 (1-5948)
QY 18 AlaAlaLeuAlaVal-----AlaProGlyProArgPheLeuValThrAla 32
DB 31 GCGCTCATGTGATGATGTTTCCAGGGTCTTCTCATGAGGCTCTCATCACCTCATCAC 90
QY 33 ProGlyIleIleArgProGlyGlyAsnValThrIleGlyValGluLeuLeu---GluHis 51
DB 91 CCTGCTGTTTGGCAGAACAGACAGAGAGCAAAATTTGGTGGAGGCCCATGAGACACT 150
QY 52 CysProSerGlnValThrValIleValAla-----GluLeuLeuValThrAlaSer 67
DB 151 ACTCCAAACAGCTGATGATGATTCATGATTTTCCAGAGAGCAAAACCTTGTTC 210
QY 68 AsnLeuThrValSerValLeuGluIleGluGlyValPheGluIleGlySerPheLeuThr 87
DB 211 CAACCCAGAGTATGATGATTCATGATTCAGAGAGGATCTTGTCACT-----CCA 261
QY 88 LeuThrLeuProSerLeuProLeuAsnSerAlaAspGluIleTyrGluLeuArgValThr 107
DB 262 ATAGAGATTCACAGAAAGAGAGTACGAGACTCCAGCAAAATCAATATGATGTGTG 321
QY 108 GlyArgThrGlnAspGluIleLeuPheSerAsnSerThrArgLeuSerPheGluThrIle 127
DB 322 CAAGTAACTGCTGCTCAAGTGAATTCGAAAAGTGTCTCTCTTCTTCAACGAGTAC 381
QY 128 ArgIleSerValPheIleGlnThrAspIleValLeuTyrIleProIleGlnIleValIle 147
DB 382 -----TTTCTGTTTCCAGACAGATTAAGCATCTATACACAGAGGTCTCCAGTAC 435
QY 148 PheArgIleValThrIleuPheSerAspPheLeuProTyrIleThrSerLeuAsnIleLeu 167
DB 436 TATCGGTTTTCATGATGATCAACACACAGACAGATGACAAACCTGTGATGTGTGAG 495
QY 168 IleIleAspProIleSerAsnLeuIleGlnIleThrLeuSerGlnIleSerAsnLeuGly 187
DB 496 TTTCAAGACTCCAGAGAGCATTTCTGTCACT-----TCTAATTCATTAACCTTAAC 546
QY 188 ValIleSerLeuThrPheGluLeuSerSerHisProIleLeuGlyAspIleSerIle--- 206
DB 547 TTCTTC---TGCGCTTACAAATTTTACAGACCTTGTCACTTGGGGAGCTTGAAGATTTG 603

```

```

QY 207 ---GlnValGlnValAsnAspGlnThrTyrTyrGlnSerPheGlnValSerGluTyrVal 225
DB 604 GCCAAATATGAACATCTCCAGAGAAATATACGTATTTTGAATCTCAGAAATATGTG 663
QY 226 LeuProIlePheGluValThrLeuGlnThrPro-----LeuTyrCysSerMetAsn 242
DB 664 TTGCCAAGCTTGAAGTCCGTCTGCACATCAGAAAGTTTTCATTTGACCGGCAT 723
QY 243 SerIleHisIleAsnGlyThrIleThrAlaIleTyrThrTyrGlyIleValProValIleGly 262
DB 724 ---GAAATTTCCAGTGTCTATCACTCAAGATCTTGATGAGAGAAAGTGAAGCT 780
QY 263 AspValThrLeuThrPheLeuProLeuSerPheTyrGlyIleValYleAsnIle----- 280
DB 781 ---GTGGCTTGTTCCTCTTGGAGTGAATAATAGATGATCTTAAAAAGATTTCCAGAC 837
QY 281 -----ThrIleThrPheIleIleAsnGlySerAlaAsnPheSerPheAsnAspGlu 298
DB 838 TCACCTCAGAGAAATCCGATATTGATGAGATGGAAAGCAACACTAAAAAGATACA 897
QY 299 MetIleAsnValMetAspSerSerAsnGlyLeuSerGluTyrIleAspLeuSerSerPro 318
DB 898 TTCCGTTCTCATTTCCAAATCTCAATGAGCTTGTGGGCACTCTGTATGATCACT--- 954
QY 319 GlyProValGluIleLeuThrThrValThrGlnSerValThrGlyIleSerArgAsnVal 338
DB 955 -----GTAAACGTATGACAGAAATACGACAGATGATATGTAGTACTGAG 999
QY 339 SerThrAsnValPhePheIleGlnHisAspTyrIleIleGluPheAspTyrThrThr 358
DB 1000 CAAGCGGATTCATATTGTGGCATCTCCATATCATGATCCATTCACAAAACCCCAA 1059
QY 359 ValLeuIleProSerLeuAsnPheThrAlaThrValIleValThrArgAlaAspGlyAsn 378
DB 1060 TATTTCAAGCCAGAGATCCATATGATGACGAGGTATGTTATCCAACTGATGTGCTCA 1119
QY 379 GlnLeuThrLeuGluGluArgAsnAsnValValIleThrValThrGlnArgAsnTyr 398
DB 1120 CCAAGT-----GCCATGTGCCAGTGTGATTCAGAGGCTTT 1155
QY 399 ThrGluTyrTyrSerGlySerAsnSerGlyAsnGlnIleValIleValIleValIle 418
DB 1156 -----CATTCATGGGAACACTTGAAGTATGGAGCTGATAGCTC 1197
QY 419 AsnTyrThrValProGlnSerGlyThrPheIleIleGluPheProIleLeuGluAspSer 438
DB 1198 ATCTGAACATACCATTTGAATGCTCA-----AGCTTACCAATCATCTTGAAGACT 1248
QY 439 SerGluLeuGlnIleValAlaTyrPheLeuGlySerIleSerSerMetAlaValHisSer 458
DB 1249 AACCATGAGACCTTCCAGAGAAACGCGAGCAACAAAGTCAATGACAGCATGACCTAC 1308
QY 459 LeuPheIleSerProSerIleThrTyrIleGlnLeuIleuThrArgAspGluAsnIleIle 478
DB 1309 CAACCCAGGAGAGATCTGGAACATCTTCTTATGATGACATTAATCTACAGAGATTAAG 1368
QY 479 ValGlySer-----ProPheGluLeuValIleSerGlyAsn-----LysArg 492
DB 1369 CCCGAGATTAATCACTGCAATTTCAATGTGAAGGCAATGCAAAATTCACCTGAGAGAG 1428
QY 493 LeuIleGluLeuSerTyrMetValValSerArgIleGlnLeuValAlaValIleGlyIle 511
DB 1429 ATCAAAATTTCCATCTCATATGATTAATAAGGAAAGATTTTCAAGTGTGGAGCA 1488
QY 512 -----GlnAsnSerThrMetPheSerLeuThrProGluAsnSerTyrThr 526
DB 1489 CCCAGAGAGATGGGCAAAATCTGTGACATGATATCTGCATATCACTCCAGATCTATC 1548
QY 527 ProIleValCysValIleValIleValTyrIleGluAspAspGlyGluIleIleSerAspVal 546
DB 1549 CTTTCTTCGCGTGTGCTTACTACCAAGTGGGAAACAGAAATTTGTGGCTGATTTCT 1608

```









Db	2401	ATC----	TTGTTGGCTGGAACCTTATGAAATGAGATCAAGAAAGCTTCTTCATTGATCTT	2457	
Qy	753	AsnLeuProTyrSerValIle	argGlyGluGluPheVal	aleGluIleThrTlePheAsn	772
Db	2458	CAAAAGCCATTTCTCAGTACGTAGTAAGAAATAGACAGAGTGAGATTCGACATTTCTGCACAAAC	2517		
Qy	773	TyrLeuLysAspAlaThrGluValValIle	IleGlyLysSerAspLysPheAspLe	792	
Db	2518	TACGTTAACGAGAGATATTATGTATGTGGAGTGGAACTGTTATACAAACCCTTC-----	2577		
Qy	793	LeuMetThrSerSerGluIle	asnAlaThrGlyHisGlnGlnThrLeuValProser	812	
Db	2572	-----TCCAGTGCCTCCACAAAGACAAAGATATCCACACAGACTTCCCATTTAAAGCC	2623		
Qy	813	GluAspGlyAlaThrValLeuPheProIle	argProThrHisLeuGly-----GluIle	830	
Db	2626	CTGTCCCTCCAGACAGTACCGTTGTATGATGTCCTATTAAGACAAAGATTCGATGATGTT	2685		
Qy	831	ProIleThrValThrAlaLeuSerProThrAlaSerAspAlaValThrGluMetIleLeu	850		
Db	2686	GAGATTAAAGCAAGTGCACAGAACCGTTGGTGCAGCGGTGTGAGAAACAACTGAAA	2745		
Qy	851	ValLysAlaGlnGlyIleGluLysSerTyrSerGlnSerIleLeuLysPheLeuThrAsp	870		
Db	2746	GTTGTACCTGAAGGGGTACAGAAATTCATTGTGACTATTGTTAACTGACCCAAAGGCA	2805		
Qy	871	AsnArgLeuGlnSerThr-----	LeuLysThrLeuSerPheSerPhe	884	
Db	2806	AAAGGAGTGGTGGAAACACAGCTAGACATGATCAAAAGCCGCAATTTAGATGACAGACTG	2865		
Qy	885	ProProAsnThrValThrGlySerGluArgValGlnIleThrAlaIleGlyAspValLeu	904		
Db	2866	CCTGACACAGAAATTGAAACCAAGATTATTCATCCAAAGGTGACCCCTGGCTCAGATTATT	2925		
Qy	905	GlyProSerIleAsnGly-----	LeuHisSerLeuIleArgMetProTyrGlyCysGly	922	
Db	2926	GAAAACCTCAATTTGATGGAAGTTAACTCAACCATCTCATTTATCATCTCTTCGGCTGGTGG	2985		
Qy	923	GluGlnAsnMetIleAsnPheAlaProAsnIle-----	TyrIleLeuAspTyrLeu	939	
Db	2986	GAGCAAAATATGATCCGATGCGATGCGCACAGCTATTATGCCACTTACTTGGACACCA	3045		
Qy	940	ThrLysLysLysGlnLeuThrAspAsnLeuLysGlyValAlaLeuSerPheMetArgGln	959		
Db	3046	GAGCAGTGGGAGACTCTCGCATTAATTCGACAACTGCAAGCTGTCACATCAAGATCGTACT	3105		
Qy	960	GlyTyrGlnArgGluLeuLeuTyrGlnArgGluAspGlySerPheSerAlaPheGlyAsn	979		
Db	3106	GGTATATGCCACAGCATGATGTGTACAAAGAACACAGTCACTTCCTTAGCGACATTTAACAAAC	3165		
Qy	980	TyrAspProSerGlySerThrTyrLeuSerAlaPheValIleuArgCysPhe---	LeuGlu	998	
Db	3166	-----CGTCATCTAGTTCTTGGCTTACAGCATATGTGCTGTAAGTCTTGGCCATGGCT	3219		
Qy	999	AlaAspProTyrIleAspIleAspGlnAsnValLeuHisArgThrTyrThrTyrPhe---	1017		
Db	3220	GCCAAAATATGTAGACAGCATATGTCATAAATCATTTGTGGAGGTGTCAGGTGGCTATTT	3279		
Qy	1018	LysGlyHisGlnLysSerAsnGlyGluPheTyrAspProGlyArgValIleHisSerGlu	1037		
Db	3280	CTGAACAGGCAACAAACAGATGGAGCGCTTCAAGAAAGATGCCCCCTGTACTTCTCGAACA	3339		
Qy	1038	LeuGlnGly-----	GlyAsnLysSerProValThrLeuMetAlaTyrIleValThr	1054	
Db	3340	ATGCAGGAGGAATTCAGAGTGTCTGGAAGAAGAAGATATATTTAACAGCTTTCATTTCTGGTT	3399		
Qy	1055	SerLeuLeuGlyTyrArg-----	LysTyrGlnProAsnIleAsp---	Val	1068
Db	3400	CGGTTGTGTGAATCCAAAACAATCTGCACATGCTATGTCCAAATGTCTGACAGGACATC	3459		
Qy	1069	GlnGlnSerIleHisPheLeuGlnSerGluPheSerArgGlyIleSerAspAsnTyrThr	1088		

Db	3460	AAGAAGCCACAATTTATTCTGAAAAGATGAGAA---	-CTGCAAAGGCCCTTACACT	3516
QY	1089	LeuAlaleu11eThrTYrAlaLeuSerSerValIglySerProlybAlaLySgluAlaLeu	1108	
Db	3517	ACAGCCCTTCACAGCCTATGCTTTGGCT---	-GCTGCAGCAACTC	3558
QY	1109	AsnMetLeuThrTrpArgAlaGluGlnGluIglyIyMetGlnPheTrpValSerSerGlu	1128	
Db	3559	AATGATGACAGGGTACTTCATGGCAGCATCAACAGAAAGGATCATTTGG---	3606	
QY	1129	SerIySLeuSerAspSerTrpGlnProArgSerLeuAspIleGluValAlaAlaTyzAla	1148	
Db	3607	-----GAAATATCAAAAGCTCACACCACCAACATTGAAGCAGCACTTCTATGCC	3654	
QY	1149	LeuLeuSerThiSphLeuGlnPheGlnThiSerGluIyIleProIleMetArgTrpLeu	1168	
Db	3655	TTGTTGGCCCTGCTGAAATAATGAAMAATTGATCAAACTGGATCCATATGTCAGATGGCTG	3714	
QY	1169	SerArgGlnArgAsnSerLeuGlyIyPheAlaSerThrGlnAspThrThrValAlaLeu	1188	
Db	3715	ACAAGATCAGAATTTTATGGGAAACATATGACAAACCAGCAACAGTATGACATTT	3774	
QY	1189	LySAlaLeuSerGluPheAlaAlaLeuMetAsnThrGluArg---ThrAsnIleGlnVal	1207	
Db	3775	CAAGCTCTCTGTAATATGAGATTCAAGTGGCTTCCCATTAAGCACTTAACTTGAATAT	3834	
QY	1208	ThrValThrGlyProSerSerProSerProLeu-----	1218	
Db	3835	ACTATTTGAACTGCAGATCGAAGTACGTATATAGTACAGAAATTAATATGAAAAATGCT	3894	
QY	1219	-----AlaValValGlnPro-----MetAlaValAsnIleSerAlaAsnGly	1232	
Db	3895	CTCCTGGCTCGAGCATGAGACCCAACTCAACCAAGACATCACTGGTGCAGCATCAAGT	3954	
QY	1233	PheGlyPheAlaIleCySglnLeuAsnValValTyrrAsnValLySAlaSerGlySerSer	1252	
Db	3955	GATGAGAAAGACAAATGACCATTTTGACATTCTTATTAACGCACGCTGG-----	4002	
QY	1253	ArgArgArgArgSerIleGlnAsnGlnGluAlaPheAspLeuAspValAlaValLySglu	1272	
Db	4003	CAGAGAAAGCAAAATGTTTGCATAAA-----TTTCATCTTAATGTTTCTGTTGAA---	4053	
QY	1273	AsnLySAspAspLeuAsnHisValAsp-----LeuAsnValCysThr	1286	
Db	4054	---AACATCCACTGAATGACATGAGAGCCAAAGGAGCCCTCATGCTCAAGATCTGCACA	4110	
QY	1287	SerPheSerGlyProGlyArgSerGlyMetAlaLeuMetGluValAsnLeuLeuSerGly	1306	
Db	4111	AGGATATCTGGAGAAAGTTGATTCTTACATGACAAATTAATATTTATTTATGCTGACGTG	4170	
QY	1307	PheMetValProSerGluAlaIleSer-----LeuSerGluThrValLyS	1321	
Db	4171	TTTTCTCCCTGATGCTGAAAGACCTTACAAAGCTTTCTTAAAGAGTGCAGACATACATCTCC	4230	
QY	1322	LySValGluTyrrAspHisGlyLySLeuAsn-----LeuTyrrLeuAspSer	1336	
Db	4231	AGATATGAGTGTGACAAATAATATGCTGCAGAAATAGACTGTTTATCATTTACTTAAACAG	4290	
QY	1337	ValAsnGluThrGln---PheCySValAsnIleProAlaValArgAsnPheLySValSer	1355	
Db	4291	GTCGCCCACTCTGAAGATGAAATGCTCCTCACATTGAATTCACAGCAATTTTGAAGTTGGC	4350	
QY	1356	AsnThrGlnAspAlaSerValSerIleValAspTyrrGlyProArgArgGlnAlaVal	1375	
Db	4351	TTTCATTACGCCAGATCAATCAAGGTGTACAGCTACTACATCTTATGATGAATAATGTAAC	4410	
QY	1376	ArgSerTyrrAsnSerGluValLySLeuSerSerCysAspLeuCySerAspValGlnGly	1395	
Db	4411	AAGTTTACCATTCAGATTAAGAAGAAACAGGCTTCTCAATTAAGATATGATATGTAACGTT	4470	
QY	1396	CysArgProCySglnAspGlyAlaSerGlySerHisHis-----	1408	
Db	4471	TGCCGATGTGACAGGAACCTGTTTCTGCTCAACCATCAGGAAGATGATGATTTCA	4530	

QY 1409 -----HisSerValIlePheIlePheCysPheIleLeu 1421  
Db 4531 TTACAAATTGAAAAAGCTGGAGACGAATGTGATTATATCTACAAACCAAGCTGCTT 4590  
QY 1422 TyrPheMetGlu 1425  
Db 4591 CGAATAGAGAA 4602

RESULT 9  
US-08-447-411-44  
Sequence 44, Application US/08447411  
Patent No. 5773243  
GENERAL INFORMATION:  
APPLICANT: FRITZINGER, DAVID C.  
APPLICANT: BREDEHORST, REINHARD  
APPLICANT: VOGEL, CARL-WILHELM  
TITLE OF INVENTION: DNA ENCODING COBRA C3, CVF1, AND CVF2  
NUMBER OF SEQUENCES: 81  
CORRESPONDENCE ADDRES:  
ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,  
ADDRESS: P.C.  
STREET: 1755 S. Jefferson Davis Highway, Suite 400  
CITY: Arlington  
STATE: Virginia  
COUNTRY: U.S.A.  
ZIP: 22202  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent in Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/447,411  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/043,747  
FILING DATE: 07-APR-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Oblon, No. 5773243man F.  
REGISTRATION NUMBER: 24,618  
REFERENCE/DOCKET NUMBER: 1126-101-0  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (703) 413-3000  
TELEFAX: (703) 413-2220  
TELEX: 248855 OPAT UR  
INFORMATION FOR SEQ ID NO: 44:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 5924 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: unknown  
TOPOLOGY: unknown  
MOLECULE TYPE: DNA (genomic)  
FEATURE:  
NAME/KEY: sig peptide  
LOCATION: 4..69  
NAME/KEY: mat peptide  
LOCATION: 70..4929  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 4..4929  
US-08-447-411-44

Alignment Scores:  
Pred. No.: 4,74e-83 Length: 5924  
Score: 885.00 Matches: 359  
Percent Similarity: 40.84% Conservative: 296  
Best Local Similarity: 22.38% Mismatches: 673  
Query Match: 12.04% Indels: 276  
DB: 1 Gaps: 56

US-10-020-095-4 (1-1428) x US-08-447-411-44 (1-5924)  
QY 18 AlaAlaLeuAlaVal-----AlaProGlyProArgPheLeuValThrAla 32  
Db 31 GCTGCTCTATGATTTGGTTTCCAGGGTCTTTCATAGGGGCTCTACACCTCATCACC 90  
QY 33 ProGlyIleIleArgProGlyGlyAsnValThrIleGlyValGluLeuLeu---GluHis 51  
Db 91 CTGCTGTTTTCGACAGACACAGAGACAAATTTGGTGAGGCCATGAGACAGT 150  
QY 52 CysProSerGlnValThrValIleVala-----GluLeuLeuSerThrAlaSer 67  
Db 151 ACTCCAAAACAGCTTACATCTTTGTCATGATTTTCCAGGAGACAAACCTTGTTTC 210  
QY 68 AsnLeuThrValSerValIleGluIleGluValIlePheGluIleGlySerPheThr 87  
Db 211 CAAACCAAGTATGATATGATTCACGACGAGGATCTGTCTACT-----CCAACT 261  
QY 88 LeuThrLeuProSerLeuProLeuAsnSerAlaAspGluIleTyrGluLeuArgValThr 107  
Db 262 ATAGAGATTCACGAAAGAGAGAGTACGACTCCAGGCAAAATCAATATGTGTTGTC 321  
QY 108 GlyArgThrGlnAspGluIleLeuPheSerAsnSerThrArgLeuSerPheGluThrIys 127  
Db 322 CAAGTAACTGCTCCAGAGTGAAGTGAAGAGTGGTCTCTCTTTCTTAACCAAGTACG 381  
QY 128 ArgIleSerValPheIleGlnThrAspIleValAlaLeuTyrIleProGlyGlnIleValIys 147  
Db 382 -----TTTCTGTTTATCCAGACATTAAGGATTTATACACAGGCTCTCCAGTAC 435  
QY 148 PheArgIleValThrLeuPheSerAspPheIysProTyrIleThrSerLeuAsnIleLeu 167  
Db 436 TATCGTGTGTTTCTATGATGATCACAACAGACAGATGAACAACTGTGATGTTGAG 495  
QY 168 IleIysAspProIysSerAsnLeuIleGlnIleThrLeuSerGlnIleSerAspLeuGly 187  
Db 496 TTTCAGATCTCCAGAAAGCATTTCTGTGCACT-----TTTAATTCAGTTGACCTTAAC 546  
QY 188 ValIleSerIleThrPheGlnLeuSerSerHisProIleGluIleAspTrpSerIle--- 206  
Db 547 TTCTTC---TGCCCTTACAAATTTTACCAACCTTGTAGTTGGGACTTGGAGGATTTG 603  
QY 207 ---GlnValGlnValAsnAspGlnThrTyrTyrGlnSerPheGlnValSerGluTyrVal 225  
Db 604 GCCAAATATGACATTCCTCCAGAGATTTATATCGCATTTTGTGATGTCAGAAATATG 663  
QY 226 LeuProIysPheGluValThrLeuGlnThrPro-----LeuTyrCysSerMetAsn 242  
Db 664 TTGCCAAGCTTTGAAAGTCCGTCGCAACATCAGAGAGTTTTATCAATTCAGCGCAAT 723  
QY 243 SerIleHisLeuAsnGlyThrIleThrAlaIysTyrThrTyrGlyIleProValIysGly 262  
Db 724 ---GAAATTTCCAGCTGCTATCACTGCAAGGATCTTGATGAGAGAGAGAGGAG 780  
QY 263 AspValThrLeuThrPheLeuProLeuSerPheTyrGlyIleIleValIleAsnIle----- 280  
Db 781 ---GTGGCTTCTCTCTTTGGAGTGAATATGATGATGCTAAAGAGTATTCAGAC 837  
QY 281 -----ThrIleThrPheIleValIleAsnGlySerAlaAsnPheSerPheAsnAspGlu 298  
Db 838 TCACATCAGCAATTCGATTTATGATGATGAGAGAGCAACTAAAGAGATTAACA 897  
QY 299 MetIleAsnValMetCysSerSerAsnGlyLeuSerGluTyrLeuAspLeuSerAspPro 318  
Db 898 TTCGGTTCGATTTCCAAATTCCAATGAGCTTGTGGCATATCTGTATGATCATC--- 954  
QY 319 GlyProValGluIleLeuThrThrValThrGluSerValThrGlyIleSerArgAsnVal 338  
Db 955 -----GTAACATGATGACAGATATGAGAGATATGATGATGATGAG 999  
QY 339 SerThrAsnValPhePheIleGlnHisAspTyrIleIleGluPhePheAspTyrThr 358  
Db 1000 CAAAGCGGATTCATATGTGGCATCTCCCTATCAGATCCACTTCAAAAACCCCAAA 1059

QY 359 ValLeuLysProSerLeuAsnPhetHraIaThrValIleValIleThrArgAlaAspGlyAsn 378  
 Db 1060 TATTTCAAGCCAGGATGCCATPATGAACTGACGGTGTATGTTACCAACCTGATGGCTCA 1119  
 QY 379 GlnLeuThrLeuGlnGluArgArgAsnValIleIleThrValIleThrGlnAsnIleTy 398  
 Db 1120 CCAGCT-----GCCCATGTGCAGTGTCCAGTGTATCAGAGCCCTT 1155  
 QY 399 ThrGluTyTrpSerGlySerAsnSerGlyAsnGlnLysMetGlnAlaValGlnLysIle 418  
 Db 1156 -----CATTCATGGGAACCATTTGATGATGAGCTGCTAAAGCTC 1197  
 QY 419 AsnTyThrValProGlnSerGlyThrPheLysIleGluPheProIleLeuGluAsnSer 438  
 Db 1198 ATCCGTAACATACCATTTGAATGCTCA-----AGCTACCATTCACCTGTAAAGCT 1248  
 QY 439 SerGluLeuGlnLeuLysAlaTyPheLeuGlySerLysSerSerMetAlaValHisSer 458  
 Db 1249 AACCATGAGACCTCCCAAGAGAACCCAGGCAACAAAGTCCATGACAGCCCATAGCCCTAC 1308  
 QY 459 LeuPheLysSerProSerLysThrTyrlleGlnLeuLysThrArgAspGluAsnIleLys 478  
 Db 1309 CAAACCCAGGAGAGATCTGGAACCTATCTTCATGTAGCCATTAATCTACAGAGATTAAAG 1368  
 QY 479 ValGlySer-----ProPheGluLeuValIleSerGlyAsn-----LysArg 492  
 Db 1369 CCCGAGATACCTTACTCTGTCATTAATTCATGGAAGGCAATGCATTCACGTAAAGCAG 1428  
 QY 493 LeuLysGlnLeuSerTyMetValValSerArgGlyGlnLeuValAlaValGlyLys--- 511  
 Db 1429 ATCAATATATTTCACTCATCTCATATTTGAATAAGGAGAAATTTTCAAGTGTGGCAGCA 1488  
 QY 512 -----GlnAsnSerThrMetPheSerLeuThrProLysAsnSerTrpThr 526  
 Db 1489 CCCAGAGAGATGGGAGATCTGTGACCATGATCAATCTCATATCACTCCAGATCTCATC 1548  
 QY 527 ProLysAlaCysValIleValIleTyrlleGlnAspAspGlyGlnIleLysSerAspVal 546  
 Db 1549 CTTCTCCCTCCGTTTGTGTGCTTACTACCAAGTGGGAAACAGCAATTTGTGTGATTTCT 1608  
 QY 547 LeuLysIleProValGln-----LeuValPheLys-----AsnLys 558  
 Db 1609 GTCTGGGTGATGTGAAGATACCTGCATGGAACCTGTGTGTGAAGAGACATCTCA 1668  
 QY 559 IleLysLeuTyTrpSerLysValIleValIleGluProSerGlyLysValSerLeuArgIle 578  
 Db 1669 ATACCAATG-----CCAGAGCTGCAATGAAATCAAAATG 1704  
 QY 579 SerValIleThrGlnProAspSerIleValIleValIleValIleValIleValIleVal 598  
 Db 1705 GAA---GGGATCTCAGGTGCTCGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1761  
 QY 599 MetAsnAlaSerAsnAspIleThrMetGluAsnValIleIleGluLeuGluLeuTyraen 618  
 Db 1762 CTCATGATGAATAATATAGATTAGCCCAAGCTTAAGATGGGACACATAGAAAGAGTAC 1821  
 QY 619 ThrGlyTyTrpLeuGlyMetPheMetAsnSerPheAlaValPheGlnGluCysGlyLeu 638  
 Db 1822 TTTGGCTGTACAGCTGGCAGTGGCCGAAATAATCTGGGTGTGTGTGTGTGTGTGTGTGT 1881  
 QY 639 TrpValLeuThrAspAlaAsnLeu---ThrLysAsp----- 649  
 Db 1882 GCTGTGACAAACGACACTTAATCTCAACACCAACAGAGATCAGCTGCAAAAGTGTCTCAG 1941  
 QY 649 ----- 649  
 Db 1942 CCTGCAATTCGAGAGCGGTGCGAGTTCTGTTTGTGCTTGTGACAGCAACGCAAGCAAGG 2001  
 QY 649 ----- 649  
 Db 2002 GGAGAAATTTCAAGATCAAGACCTGCGTAATGCTGTGAAGATGTCAATGATGAGAAACCC 2061

QY 649 ----- 649  
 Db 2062 ATGGGGTACACTTGTGAAGAGGTGCAAAATACATCCAGAGAGAGATGCTTGAAGCT 2121  
 QY 650 -----TyrlleAspGlyValTyraen---AsnAlaGluTyraIa 661  
 Db 2122 GCCCTTCCTGAATGCTGTGCTACATCAAGGGGGTCCGAGTGAAGAAACCAAGGAGAGC 2181  
 QY 662 GluArgPheMet-----GlnGluAsnGlu---GlyHisIleValAspIleHisAspPhe 678  
 Db 2182 GAGTGTTCCTGGCAAGAGATGATTAATGAAGTGGTTTCAATGCAAGATGTGATATCATC 2241  
 QY 679 SerLeuGlySerSerProHisValArgLysHisPheProGluThrTyrlleTrpLeu--- 697  
 Db 2242 TCA-----AGTGTGATTTCCCAAGAGTTGTGTGGCTTAACA 2280  
 QY 698 -----AspThrAsnMetGlyTyraGlyIleTyrlleGlnLysPheGluValThr 712  
 Db 2281 AAGCACTTGACCGAGAGCCTTAACAGTCAAGGATTTCAAGCAACAAATGCTTTTAT 2340  
 QY 713 ValProAspSerIleThrSerTrpValAlaThrGlyPheValIleSerGluAspLeuGly 732  
 Db 2341 CTGAGGATTCATCAACAACCTGGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2400  
 QY 733 LeuGlyLeuThrThrProValGluLeuGlnAlaPheGlnProPhePheIlePheLeu 752  
 Db 2401 ATC---TGCTGGCTGAACCTTATGATTAATTAAGATCAATGAAGTCTTCTTCATGATCT 2457  
 QY 753 AsnLeuProTySerValIleArgGlyGlnGluPheAlaLeuGluIleThrIlePheAsn 772  
 Db 2458 CAAAGCCATATTCAGTGTGTAAGATGAGCAGGTGAGATTCGAGCTATTCGACAAAC 2517  
 QY 773 TyrlleLysAspAlaThrGluValIleValIleIleGluLysSerAspLysPheAspIle 792  
 Db 2518 TACGTTAACGAGATATTATATGTGAGGTGAACCTGTATACAAACCCCTTC----- 2571  
 QY 793 LeuMetThrSerSerGluIleAsnAlaThrGlyHisGlnGlnThrLeuLeuValProSer 812  
 Db 2572 -----TGCAAGTCTCCCAAAAGACAAATATCCGACACACAGTCCCAATTAAGCC 2625  
 QY 813 GluAspGlyAlaThrValLeuPheProIleArgProThrHisLeuGly-----GluIle 830  
 Db 2626 CTGTCCCTCAGAGAGTACCGTTGTGATAGTCCCATTAAGACCAAGATGATGATGTT 2685  
 QY 831 ProIleThrValIleThrAlaLeuSerProThrAlaSerAspAlaValIleThrGlnMetIleLeu 850  
 Db 2686 GAGATTAAAGCAAGAGTCCAGAGAACCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2745  
 QY 851 ValLysAlaGluGlyIleGluLysSerTySerGlnSerIleLeuLeuAspLeuThrAsp 870  
 Db 2746 GTTGTACCTGAAGGGGTACAGAAATCATTTGTGACTATTTGTAATCTGAGCCCAAGGGA 2805  
 QY 871 AsnArgLeuGlnSerThr-----LeuLysThrLysSerPheSerPhe 884  
 Db 2806 AAAGAGTTGTGGAACAACAGCTAGAGATCAAGCCCGCAAAATTTAGATACAGAGTG 2865  
 QY 885 ProProAsnThrValIleThrGlySerGluArgValGlnIleThrAlaIleLysAspValLeu 904  
 Db 2866 CTGACACAGAAATGAACCAAGATTATCAACCAAGTGAACCTGTGTGCTCAGATTAAT 2925  
 QY 905 GlyProSerIleAsnGly-----LeuAlaSerLeuIleArgMetProGlyGlyCysGly 922  
 Db 2926 GAAAACTCAATGATGAAGTAAGTAACTCAACATCTCATATTAATCACTCTTGTGGCTGGG 2985  
 QY 923 GluGlnAsnMetIleAsnPheAlaProAsnIle-----TyrlleLeuAspTyrlleu 939  
 Db 2986 GAGCAAAATATGATCCGATGCGCAGCAGCAAGTATTGCACTTAATCTGACCAACCAACA 3045  
 QY 940 ThrLysLysLysGlnLeuThrAspAsnLeuLysGlyLysAlaLeuSerPheMetArgGln 959  
 Db 3046 GAGCAGTGGAGACTCTGGCAATTAATCGCAGAGATGTCAATGATGATGCTGACT 3105  
 QY 960 GlyTyrlleArgGluLeuLeuTyrlleGlnArgGlySerPheSerAlaPheGlyAsn 979

```

Db 3106 GGTATCCAGCAGATGCTGATCAAGAAAGCATATCTCTATGACGATTACAAAC 3165
Qy 980 TyrAspProSerGlySerThrTrpLeuSerAlaPheValLeuAlaGlySerPhe---LeuGlu 998
Db 3166 -----CGTGCATCTAGTCTTGGCTAACACACATATGCGTAAGAAGCTTTGCCATGGCT 3219
Qy 999 AlaAspProTyrIleAspIleAspGlnAsnValLeuHisArgThrTyrTrpIleu--- 1017
Db 3220 GCCAAATGATGACGACATATGATCATGAAATCATTTGTGGAGGTGAGAGGTGCTGAT 3279
Qy 1018 LysGlyHisGlnLysSerAsnGlyGluPheTrpAspProGlyArgValIleHisSerGlu 1037
Db 3280 CTGAACAGGCAACACAGATGAGCGCTCAAGAAATGCCCTGACATCTTCGAAACA 3339
Qy 1038 LeuGlnGly-----GlyAsnLysSerProValThrLeuThrAlaTyrIleValThr 1054
Db 3340 ATGCAGGAGAGAAATTCAGAGTGTGAAGAGAAAGATATTTAAACAGCTTTCATCTGGTT 3399
Qy 1055 SerLeuLeuGlyTyrArg-----LysTyrGlnProAsnIleAsp-----Val 1068
Db 3400 GCGTGTGATGATCCAAACATCTGCATGACTATGTCATAGTCAAGCAGCAGATC 3459
Qy 1069 GlnGlnSerIleHisPheLeuGlnSerGluPheSerArgIylIleSerAspAsnTyrThr 1088
Db 3460 AAGAAGGCCCAAAATATTATCTCAAAAAGATGAGAAA---CTGCAAGGCGCTTACACT 3516
Qy 1089 LeuAlaLeuIleThrTyrAlaLeuSerSerValGlySerProValAlaLysGlnAlaLeu 1108
Db 3517 ACAGCCCTCACAGCCTTATGCTTGGCT-----GCTGCAGACCAACTC 3558
Qy 1109 AsnMetLeuThrTyrArgAlaGlnGlnGlyGlyMetGlnPheTrpValSerSerGlu 1128
Db 3559 AATGATGACAGGATCTCATGCGACGATCAACAGAGAGGATCATCTGG----- 3606
Qy 1129 SerLysLeuSerAspSerTrpGlnProArgSerLeuAspIleGluValAlaIleTyrAla 1148
Db 3607 -----GAAGATACAAAGCTCACACCAACCAACATGTAAGGCACTTCTATGGCC 3654
Qy 1149 LeuLeuSerHisPheLeuGlnPheGlnThrSerGlyIleProIleMetArgTrpLeu 1168
Db 3655 TTGTTGGCGCTCGAAATGAGAAATGATGATCAACTGCTCCATAGTACAGTGGTG 3714
Qy 1169 SerArgGlnArgAsnSerLeuGlyGlyPheAlaSerThrGlnAspThrThrValAlaLeu 1188
Db 3715 ACAATGACAGATTTTATGGGAAACATATGACAAACCCAGACAGATGATGCAATT 3774
Qy 1189 LysAlaLeuSerGlnPheAlaAlaLeuMetAsnThrGluArg---ThrAsnIleGlnVal 1207
Db 3775 CAAGCTCTTGCTGAATGATGATTCAGATGCTTACCCATAGAGACTTAACCTTAGATAT 3834
Qy 1208 ThrValThrGlyProSerSerProSerProLeu----- 1218
Db 3835 ACTATTGAACCTGCAGATCGAAGAGTACTTAAAGTACAGAAATTAATTGAAATGCT 3894
Qy 1219 -----AlaValAlaGlnPro-----MetAlaValAsnIleSerAlaAsnGly 1232
Db 3895 CTCCTGGCTCGGACAGTACAGCAAACTCAACAGACATCACTGTACAGCATCAGCT 3954
Qy 1233 PheGlyPheAlaIleCysGlnLeuAsnValValTyrAsnValValLysSerGlySerSer 1252
Db 3955 GATGGAAGAACACATGACATTTTGACATTTATTAACGACAGCTTG----- 4002
Qy 1253 ArgArgArgArgSerIleGlnAsnGlnGluAlaPheAspLeuAspValAlaValLysGlu 1272
Db 4003 CAGAGAGAGCAAAATGTTGCAATTA---TTTCATCTTAATGTTTCTGTTGAA--- 4053
Qy 1273 AsnLysAspAspLeuAsnHisValAsp-----LeuAsnValCysThr 1286
Db 4054 ---AACATCCACTTGATGACGAGGAGCCCAAGGAGCCCTCATGCTTAAGATCTGCACA 4110
Qy 1287 SerPheSerGlyProGlyArgSerGlyMetAlaLeuMetGluValAsnLeuLeuSerGly 1306

```

```

Db 4111 AGGTATCTGGAGAGTGTGATCTTCAATGACATATATGATTTCTATGCTGACTGCT 4170
Qy 1307 PheMetValProSerGlnAlaIleSer-----LeuSerGlnThrValLys 1321
Db 4171 TTCTCCCGTACGCTGTAAGACCTTACAGGCTTTCTAAAGAGTGGACAGATCATCTCC 4230
Qy 1322 LysValGlnTyrAspHisGlyLysLeuAsn-----LeuTyrLeuAspSer 1336
Db 4231 AGATATCAAGTTCACATATATATGCTGCAAGAAAGTACGCTGTATCTTATCTTAAC 4290
Qy 1337 ValAsnGlnThrGln---PheCysValAsnIleProAlaValArgAsnPheLysValSer 1355
Db 4291 GTCTCCACTCTTAAGATGATGATCTGCACTTTAAAGATTTCTCAAGCATTTTGAAGTTGGC 4350
Qy 1356 AsnThrGlnAspAlaSerValSerIleValAspTyrTyrGlnProArgArgGlnAlaVal 1375
Db 4351 TTGATTCAGCCAGAGATGATGACAGGTGACAGTACTACAAATCTAGTAAATGTAACC 4410
Qy 1376 ArgSerTyrAsnSerGlnValLysLeuSerSerCysAspLeuCysSerAspValGlnGly 1395
Db 4411 AAGTTCATCCATCCAGTAAAGAGACAGGCTTCTCATATGATATGATATGTTAAAGTT 4470
Qy 1396 CysArgProCysGlnAspGlyAlaSerGlySerHis----- 1408
Db 4471 TGCCGATGTGACAGAGAAACCTGCTCTCGTCAACCATAGAGAAAGATTGATTTCCA 4530
Qy 1409 -----HisSerSerValIlePheIlePheCysPheLysLeu 1421
Db 4531 TTCAAAATGAAAAGCTGCGAGACGAAATGATGATATGTCATCAAAACCAAGCTGCTT 4590
Qy 1422 TyrPheMetGlu 1425
Db 4591 CGAATGAGAGAA 4602

RESULT 10
US-08-447-411-75
; Sequence 75, Application US/08447411
; Patent No. 5773243
; GENERAL INFORMATION:
; APPLICANT: FRITZINGER, DAVID C.
; APPLICANT: BREDEHORST, REINHARD
; APPLICANT: VOGEL, CARL-WILHELM
; TITLE OF INVENTION: DNA ENCODING COBRA C3, CVF1, AND CVF2
; NUMBER OF SEQUENCES: 81
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBION, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
; ADDRESS: P. C.
; STREET: 1755 S. Jefferson Davis Highway, Suite 400
; CITY: Arlington
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/447,411
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/043,747
; FILING DATE: 07-APR-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Obion, No. 5773243man F.
; REGISTRATION NUMBER: 24,618
; REFERENCE/DOCKET NUMBER: 1126-101-0
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 413-3000
; TELEFAX: (703) 413-2220
; TELEX: 248855 OPAT UR
; INFORMATION FOR SEQ ID NO: 75:

```



```

/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 4138 base pairs
/ TYPE: nucleic acid
/ STRANDEDNESS: unknown
/ TOPOLOGY: unknown
/ MOLECULE TYPE: DNA (genomic)
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: 3..4001
US-08-447-411-75

Alignment Scores:
Pred. No.: 6,68e-71 Length: 4138
Score: 768.50 Matches: 306
Percent Similarity: 41.88 Conservative: 212
Best Local Similarity: 24.74 Mismatches: 465
Query Match: 10.46 Indels: 255
DB: 1 Gaps: 46

US-10-020-095-4 (1-1428) x US-08-447-411-75 (1-4138)

QY 349 TyrIleIleGluPhePheAspTyrThrValLeuLeuProSerLeuAsnPheThrAla 368
DB 66 TATCAGATCTACTTCAAAAACCCCAAAATATTCACAGCCAGAAATGCCATATGAATCG 125
QY 369 ThrValIleValThrArgAlaAspGlyAsnGlnLeuThrLeuGlnGluArgAsnAsn 388
DB 126 ACGGTATAGTTACCAAACTGATGCTCACCAAGCTGCC----- 164
QY 389 ValValIleThrValThrGlnArgAsnTyrThrGluTyrTyrSerGlySerAsnSerGly 408
DB 165 CATGCCAGTGTATCAGAGCCATTCATCTGAG-----GGA 203
QY 409 AsnGlnIleMetGluAlaValGlnIleAsnTyrThrValProGlnSerGlyThrPhe 428
DB 204 ACCACTTGTAGTGGAGCTGCTAAGCTCTTCCTGACACACACCAAAATGCTCAA--- 260
QY 429 IysIleGluPheProIleLeuGluAspSerSerGluLeuGlnIleLysAlaTyrPheLeu 448
DB 261 -----AGCTACCGATCTCTGTAGAACATACATGAGACCTCCCAAGAGACGCCAG 314
QY 449 GlySerIleSerSerMetAlaValHisSerLeuPheLysSerProSerIleThrTyrIle 468
DB 315 GCATTAAGTCCATGACGACCAAGCTTACCAAAACCAGAGAGATCTGAAAATATCTT 374
QY 469 GlnLeuLysThrArgAspGluAsnIleLysValGlySer-----ProPheGluLeuVal 486
DB 375 CATGTAGCCATTACATCTACAGAGATTAAAGCCGGAGATTAACCTGCTCAATTTCAAT 434
QY 487 ValSerGlyAsn-----LysArgLeuLysGluLeuSerTyrMetValIleValSer 502
DB 435 GTGAGGGGCAATGCAAAATTCACCTGAACCGATCAAAATATTTACATACCTCACTAGAA 494
QY 503 ArgGlyGlnLeuValAlaValGlyLysGlnAsnSer-----Thr 515
DB 495 AAAGGGAAGATTTTCAAGGTGGCAGGCAACAGAGGAGATGGGAGAAATCGGTGAC 554
QY 516 Met---PheSerLeuThrProGluAsnSerTyrThrProLysAlaCysValIleValTyr 534
DB 555 ATGAATCTACATATCACTCCAGAT-----CTCATTCCTTCTCCGCTTGTGGCTTAC 608
QY 535 TyrIleGluAspAspGlyIleIleSerAspValLeuLysIleProValGln----- 552
DB 609 TACCAAGTGGAAACATGAATAATGTGGCTGATGTCTGGGTGATGTGAAGATACC 668
QY 553 -----LeuValPheLys-----AsnLysIleLysLeuTyr 562
DB 669 TGCATGGGAACGTGGTTGTGAAGAAGGAGCACTTCCAGAGACATGATGATCAATG--- 725
QY 563 TrpSerIleValIleAlaGluProSerGluLysValSerLeuArgIleSerValThrGln 582
DB 726 -----CCAGAGAGCTGCAATGAATAATCAAAATTGAA---GGGGAT 761

```

```

QY 583 ProAspSerIleValGlyIleValAlaValAspLysSerValAsnLeuMetAsnAlaSer 602
DB 762 CCAAGTGTGGATTTGGCTTGTGCTGTGACCAAGAGAAATATGTTCTCATATGATATA 821
QY 603 AsnAspIleThrMetGluAsnValAlaHisGluLeuGluLeuTyrAsnThrGlyTyrTyr 622
DB 822 TATTAAGATTACCAAGCTAATAGATATGGGACACAAATGAAAAGAGTCACTTGGCTGACA 881
QY 623 LeuGlyMetPheMetAsnSerPheAlaValIleGlnGlnIleCysGlyLeuThrValLeuThr 642
DB 882 GCTGCACTGGCCAGAAATATCTGGGTGTGTGAAGATGCTGAGCTGGCTGTGCAACC 941
QY 643 AspAlaAsnLeuThr----- 647
DB 942 AGCACTATCTCAACACCAAAAGAGATCACTGCAAAAGTGTCTCAAGCTTCAATGGG 1001
QY 647 ----- 647
DB 1002 AGCGCTCGCAGTTCTGTTTCTGCTGACAGCAACGCAAGCAAGCGCACAGTTTCAG 1061
QY 648 -----LysAspTyrIleAspGlyValTyrAspAsn----- 657
DB 1062 GATCAAGACTGCGTAAATGCTGTGAAGATGCAATGAGAACCCCATGGGCGACACT 1121
QY 658 AlaGluTyrAlaGluArgPheMetGluGluAsnGluGly----- 670
DB 1122 TGTGAAAAGCGTGAATAATACATCCAGAGGGAGATGCTTGAAGCTGCTCTCGAA 1181
QY 671 -----HisIleValAspIleHisAsp----- 677
DB 1182 TGTCTGTACTATCAATCAAGGATCCAGATATCAATAAACCAGAGAGGAGTGTGTTCTG 1241
QY 678 -----PheSerLeuGlySerSerProHisValArgLys 688
DB 1242 GCAGAGAGTGAATTTGAAGATGATTTATTTGGAGAGGTAAACATCACTCA---AGGCT 1298
QY 689 HisPheProGluThrTrpIleTrpLeu-----AspThr 699
DB 1299 GATTTCTGTGAGAGTGTGTGTGTGCTATGAGAGCGTGTCTGAACCTTCAACGTTAA 1358
QY 700 AsnMetGlyTyrArgIleTyrGlnGluPheGluValThrValProAspSerIleThrSer 719
DB 1359 GCGATTTCAGACAGATGTACCTTTTAT-----CTGAGGAGATTCATCACAACC 1409
QY 720 TrpValAlaThrGlyPheValIleSerGluAspLeuGlyLeuThrThrThr--- 738
DB 1410 TGG-----GAGTTGCTGGCTGGGCGCTTTCACCCACCAAA 1445
QY 739 -----ProValGluLeuGlnAlaPheGlnProPhePheIlePheLeu 752
DB 1446 GGGATCTGTGTGGCTGAACCTTATGAATAACATGATGAAAGACTTTCATTTGATCTT 1505
QY 753 AsnLeuProTyrSerValIleArgGlyGluGluPheAlaLeuGluIleThrIlePheAsn 772
DB 1506 CAATGCGGTATTCAGTATGTAAGATGACAGGAGGAAATTCGAGCTGTTTGTACAAAC 1565
QY 773 TyrLeuLysAspAlaThrGluValIleValIleIleGluLysSerAspLysPheAspIle 792
DB 1566 TACGCTGCAAGATATTTATGATGAGATGGAACCTGTATACAGCCCAAGCTTTCGACGT 1625
QY 793 LeuMetThrSerSerGluIleAsnAlaThrGlyHisGlnGlnThrLeuLeuValProSer 812
DB 1626 GCTTCCACAGAAAGTCAA-----AGATACCAAGACAGTTCGCAATTAAGCC 1673
QY 813 GluAspGlyAlaThrValLeuPheProIleArgProThrHisLeuGly-----GluIle 830
DB 1674 CTGTCTCCAGAGGAGTATCTGTGTATGATGCCATTAAGCAAGAGATGATGATGTT 1733
QY 831 ProIleThrValThrAlaLeuSerProThrAlaSerAspAlaValThrGlnMetLeu 850
DB 1734 GAGGTATACAGAAAGTGTCCAGAGAGTGTGATGTGTGATGTGTGAAGAAATCTGAA 1793
QY 851 ValLysAlaGluGlyIleGluLysSerTyrSerGlnSerIleLeuLeuAspLeuThrAsp 870

```

Db	1794	GTGTGACCTGAAGGGGAAATGCAAAAGTATTGTTACTATTATTGAATGACACCATCA	1853
Oy	871	AsnArgLeuGlnSerThr-----LeuTyThrLeuSerPheSerPhe	884
Db	1854	AAAGGAATTGGTGAACAACAGGTAGAAATTGGTCAAAAGCAATTAATTAAATGACAGGGTT	1913
Oy	885	ProProabnThrValThylGlySerGluArgValGlnIleThrAlaIleGlyAspValLeu	904
Db	1914	CCTGATACGGAATGTGAAACCAAGATTACTATTCAAGGTGATCTCGGTGCTCAGACTATT	1973
Oy	905	GlyProSerIleAsnGly-----LeuAlaSerLeuIleArgMetProTyGlyCysGly	922
Db	1974	GAAACTCAATTGATGGAAGTAACTCAACCATCTCATTTATGACTCTTTTGGCTGTGG	2033
Oy	923	GluGlnAsnMetIleAsnPhe---AlaProAsnIle-----TyrlleuAspTryLeu	939
Db	2034	GAGCAAAATATGATCCGCGATGACTGCACACAGTTATTGGCACTTACTGACCGACCA	2092
Oy	940	ThrIlyblyblyGlnLeuThrAspAsnLeuTyblyblyAlaLeuSerPheMetArgGln	959
Db	2094	CAGCAGTGGAGACTCTCGCGATTAATTCGCGAGCTGTAAGCTGTCACTGATCACTGAT	2153
Oy	960	GlyTyArgIlnArgGlnLeuLeuTyArgIlnArgIlnAspGlySerPheSerAlaPheGlyAsn	979
Db	2154	GGTATATGCCACGACGTGGTGTGACAGAAAGCAGACCACTTCCTATGACGACATTTCAAA	2213
Oy	980	TyrAspProSerGlySerThrTryLeuSerAlaPheValLeuArgCysAsp---LeuGlu	998
Db	2214	-----AATGCATCTAGTTCTTGGCTTAACAGCATATGTTGTAAATCTTTGGCTTGGCT	2266
Oy	999	AlaAspProTyrlleAspIleAspGlnAsnValLeuHlsArgThrTyThrTryLeu--	1017
Db	2268	GCCAAATATGTAAAGACATTAAACATGAATTCGTTTGGAGTATAGAGGGCTGTGATT	2322
Oy	1018	LysGlyHlsGlnLysSerAsnGlyGlnPheTryAspProGlyArgValIleHisSerGlu	1037
Db	2328	CTGAAACAGGCACAGCAAGATGAGATGTTCAGAGAAAAGCCCTGTACTTTTGGAAACA	2387
Oy	1038	LeuGlnGlyGlyAsnLysSerProVal-----ThrLeuThrAlaTyrlleValThr	1054
Db	2388	ATGCAGGGAGGCATTCAAGGTGCTGAACACAGAGATCTTTTAACAGCTTTCATTCTGGTT	2447
Oy	1055	SerLeuLeuGlyTyArgLysTyGlnProAsnIleAspValGlnGluSerIleHisPhe	1074
Db	2448	GCCTGTGTGGATCCAGATCAATC-----TGCATGATATATCAATCAATTT	2492
Oy	1075	LeuGlnSerGluPheSerArgGly-----IleSer	1084
Db	2493	CTAAGCACACAGCATCATGAAGGCCACACAGTTATTATCTCAAAAAGTATGACAAATGCA	2552
Oy	1085	AspAsnTyThrIleValLeuLeuIleThrTyrlaLeuSerSerValGlySerProLysAla	1104
Db	2553	AGGCTTACATCAACAGCCCTCAACACCTATGCTTTGGCT-----GCT	2599
Oy	1105	LysGluAlaLeuAsn-Met-----LeuThrTryArgAlaGluGlnGlyGlyMetGln	1122
Db	2595	GCAGAGACACATCAATGAATGACAGGGTACATCGACGATCAACAGAGAAAGGA-----	2646
Oy	1122	nPheTryValSerSerGluSerLysLeuSerAspSerTryGlnProArgSerLeuAspIle	1142
Db	2647	-----ATCGTTGGGAAGAACTTAACGCCCAACCCATAAC-AT	2683
Oy	1142	eglValAlaAlaTyrlaLeuLeuSerHisPheLeuGlnPheGlnThrSerGluGlyIle	1162
Db	2684	TGAAGGACCTTCTATGCTTGTGTCCTGCTGTGAATAATGAAGAAATTTGTGAGCGCG	2743
Oy	1162	eProIleMetArgTryLeuSerArgGlnArgAsnSerLeuGlyGlyPheAlaSerThrGln	1182
Db	2744	TCTGTAGTCCAAATGCTGATGATCAGCAATATATATGAGGGGAATCATGACAAACCA	2803
Oy	1182	nAspThrThrValAlaLeuLysValaLeuSerGluPheAlaAlaLeuMetAsnThrGluArg	1202

```

Db      2804  AGCAACAGTTATGATGTTTCAAGCTCTTGCGATATATGAGATTAGAGTTGCTTACCATTAA 2863
Qy      1202  g---ThraenilegnValthrValthrGlyProSerSerProSerPro----- 1217
          :      |||      :      :      :      :      :      :      :      :
Db      2864  GGACTTAACTTAACTGATATTACTATTGAACCTGCGATGACGAGATGACATTAAGTATCAG 2923
Qy      1218  -----LeuValValGlnProMetLeuValAsn----- 1227
          :      :      :      :      :      :      :      :      :
Db      2924  AATTAAATTGAAAAATGCTCTTCGCTGCTCAGACAGTGAAGACCAAACTCAACGAAAGCTT 2983
Qy      1228  ----IleSerIleAsnGlyPheGlyIlePheAlaIleCysGlnIleuAsnValValIlyrAsnVa 1246
          :      :      :      :      :      :      :      :      :
Db      2984  CACGTGTGCACACATCAGGTGATGGAAGAAAGCAACATGACCATTTTGAACGGTCTTAATTC 3043
Qy      1246  llyeAlaSerGlySerSerArgArgArgSerIleGlnAsnGlnAlaPheAspLe 1266
          :      :      :      :      :      :      :      :      :
Db      3044  ACAATTG-----AGGGAGGATGCAAAATGTTTGCAACAA-----TTCCTCTCT 3085
Qy      1266  waSpValAlaVal-----LyGlnAsnIlyrAspLeuAsnI 1279
          :      :      :      :      :      :      :      :      :
Db      3086  TGAATGTTCTGTGTAAGAAAGCTCCAGTTGAAGTTAAAGAGCAAGGAGCAAGGAGAC 3145
Qy      1279  sValAspLeuAsnValCysThrSerPheSerGlyProGlyIlyrSerGlyMetAlaLeu 1299
          :      :      :      :      :      :      :      :      :
Db      3146  CCTCAAGTCAAAATCTGCACACTAGTATCTGGAGAACTGATCTTCATATGACATTAAT 3205
Qy      1299  tGlnValAsnIleuSerGlyPheMetValProSerGlnAlaIleSer---LeuSerG 1318
          :      :      :      :      :      :      :      :      :
Db      3206  TGAATGTTCTTAAGTCTGACTGGTTTGTGTCCTGATACATGAAGACCTTACGAGGCTTTCTAA 3265
Qy      1318  uThrValIlyrAsp-----ValGlnIlyrAspHis-----GlyIly 1329
          :      :      :      :      :      :      :      :      :
Db      3266  AGGAGTCCACAGATATATCTCCATGTTGAAATTAAACAATTAATATGCTCAGAAAGAAAC 3325
Qy      1329  sLeuAsnIleuIlyrIleuAspSerValAsnGluThrGln---PheCysValAsnIleProAl 1348
          :      :      :      :      :      :      :      :      :
Db      3326  TGTATATCTATTACTTACATGACAAGGCTCTCCCACTCGAAGATGAATGCTGCACCTTAAGT 3385
Qy      1348  aValArgAsnIlyrAsnIleuSerAsnThrGlnAspAlaSerValSerIleValAspIlyrTy 1368
          :      :      :      :      :      :      :      :      :
Db      3386  TCTCAGACATTTTGAAGTGGCTTCATTTACAGCCAGATTCAGTCAAGGGTTCACACTACTA 3445
Qy      1368  rGlnProArgArgGlnAlaValArgSerIlyrAsnSerGlnValIlyrLeuSerSerCysAs 1388
          :      :      :      :      :      :      :      :      :
Db      3446  CAATCTAGATGAAAAAATGTATACCAAGATCTACCATTCAGATGAAGTAAAGCAAGGCTTCTCA 3505
Qy      1388  pLeuCySerAspValGlnGlyCysArgProCysGlnAspGlyIlyrAsn 1404
          :      :      :      :      :      :      :      :      :
Db      3506  TAAATATGCTTGTTAGCTTCTTCCCATGTCGACGAAGAAACCTGTTCC 3554

RESULT 11
US-08-662-227-33
; Sequence 33, Application US/08662227
; Patent No. 5922320
;
; GENERAL INFORMATION:
;
; APPLICANT: VOGEL, CARL-WILHELM
; APPLICANT: BREDEHORST, REINHORST
; APPLICANT: KOCK, MICHAEL
; APPLICANT: FRITZINGER, DAVID
; TITLE OF INVENTION: RECOMBINANT PROCVF
; NUMBER OF SEQUENCES: 39
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
; ADDRESS: P. C.
; STREET: 1755 S. JEFFERSON DAVIS HIGHWAY
; CITY: ARLINGTON
; STATE: VA
; COUNTRY: USA
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS

```



QY	813	GLUASRGYLAATHValLeuPheProLeaGProThrhHisLeuGly-----GluLe	830
Db	1674	CTGTCCTCCAGGGCAGATCCGTTTGATAGTCCATTAGACCAAGGATTCGATATGTT	1733
QY	831	ProIleThrValIleThAlaLeuSerProThrhAlaSerAspAlaValThGlnMetIleLeu	850
Db	1734	GAGGTTCACAGCAAGCTCCAGGAGACAGTTGATGTCAGATGGGTGTGAAGAAGAACTGAA	1793
QY	851	VAllybAlaGluGlyIleGluLeuSerTySerGlnSerIleLeuLeuAspLeuThraSp	870
Db	1794	GTTGACTCCGAAGGGGAATGGAAGAAAGTATGTTGTTACTATTATTAAGTACGCCACATACA	1853
QY	871	AsnArgLeuGlnSerThr-----LeuYsrThLeuSerPheSerIhe	884
Db	1854	AAAGCAATTTGGTGGAACAGGTAGCAATTTGCTCAAAAGCCAAATTAATTAATGACAGGGTT	1913
QY	885	ProProAsnThrValThGlySerGluArgValGlnIleThrhAlaIleGlyAspValLeu	904
Db	1914	CTGTATACGGAATGAATGAACCAAGATTACTATTCAGAGTGATCCGTGGTCCAGACTATT	1973
QY	905	GlyProSerIleAsnGly-----LeuAlaSerLeuIleArgMetProTyGlyCyGly	922
Db	1974	GAAAACTCAATTGATGAGGAATGAATCAACCATCTATTATCACCTCTTTGGCTGTGGG	2033
QY	923	GluGlnAsnMetIleAsnPhe---AlaProAsnIle-----TyrlleLeuAspTyrlleu	939
Db	2034	GAGCAAAATATGATCCGATGACGTGCACACAGTTATTCACCTTACTTACCTGACACCA	2093
QY	940	ThrlsYsIbYsGlnLeuThraAspAsnLeuYsGluYsAlaLeuSerPheMetArgIln	959
Db	2094	CAGCAGTGGGAGACTCTCGGCATTAATCGCAGACTGAAGCTGTCAATCCAGATCACTACT	2155
QY	960	GlyTyrglnArgGluLeuLeuTyrglnArgGluAspGlySerPheSerAlaPheGlyAsn	979
Db	2154	GATTATGCCACAGACTGTGTGATCAAGAAAGACACCAATTCCTATGACAGCATTTACAAC	2213
QY	980	TyrAspProSerGlySerThrTrpLeuSerAlaPheValLeuArgCyPhe---LeuGlu	998
Db	2214	-----ACGTGCATCTAGTTCTTGGCTTAACAGCATATGTTGTAAATCTTTGGCTTGCT	2267
QY	999	AlaAspProTyrlleAspIleAspIlnAsnValLeuHisArgThyThrTrpLeu---	1017
Db	2268	GCCAAATTTGTAAAGACATTAAACCATGAATAATCGTTGTGAGGTATGAGGCGCTGATT	2327
QY	1018	LysGlyHisGlnLeuSerAsnGlyGluIbPheTrpAspProGlyArgValIleHisSerGlu	1037
Db	2328	CTGAACAGACCAACGACAGATGAGATGTTCCAGAAAGAACGCCCTGTACTTTTGGAACA	2387
QY	1038	LeuGlnGlyGlyAsnLeuSerProVal-----ThrLeuThrhAlaTyrlleValThr	1054
Db	2388	ATGCAGGAGGACATTCAAGTGCTGAACCGAAGAGATCTTTTAACAGCTTTCATTTGGTT	2447
QY	1055	SerLeuLeuGlyTyArgIbTyrcIlnProAsnIleAspValGlnGlnSerIleHisPhe	1074
Db	2448	GCGTGTGGGAATCCAGATCAATC-----TGCATATGATATATCAATATT	2492
QY	1075	LeuGlnSerGluPheSerArgGly-----IleSer	1084
Db	2493	CTAGACAGACAGATACGTAAAGCCACACATTTATTACTCAAAAGTATGAGAAACTGAA	2555
QY	1085	AspAsnTyrrThrLeuAlaLeuIleThrTyrlaLeuSerSerValGlySerProIbAsla	1104
Db	2553	AGGCTTACACTACAGCCCTCACACGCTATGCTTTGGCT-----GCT	2594
QY	1105	LysGluAlaLeuAsn-Met-----LeuThrTrpArgAlaGluGlnGlyGlyMetGcl	1122
Db	2595	GCAGAACCACTCAATGATGACAGGCTACTCAGGACGATCAACACAGAAAGA-----	2648
QY	1122	nPheTrpValSerSerGlnSerTyLeuSerAspSerTrpGlnProArgSerLeuAspIln	1142
Db	2647	-----ATCGTTGGAGAGAACTTAACGCCCAACCCCATATAC---AT	2683

```

QY      1142  eglValAlaIaIaValAlaLeuLeuSerH1.sphLeuG1.Pheh1ThrSerGluGly1 1162
Db      2684  TGAAGGCACTTCCTATGCTTGTGTGGCCCTGCTGAAAATGAAGAAATTTGTTGAGCGCG 2743
QY      1162  eProlIewerArgTrrPleuSerAArgInAArgAenSerLeuGlyCylPheAlaSerThrG1 1182
Db      2744  TCTGTATGTCAAATGGCTGATAGATACAGAAATATTATGGGGAAACATATGAGACAAACCA 2803
QY      1182  nasPThrValAlaLeuValaLeuSerGluPheAlaIaLeuMetAnthrGluAr 1202
Db      2804  AGCAACAGTATGATGATGTTTCMAAGCTCTTGCTGAAATATGAGATTCAGATGCTTACCCATA 2863
QY      1202  g---Thrasn1IegInValThrValThrGlyProSerSerProSerPro----- 1217
Db      2864  GGACTTAACTTAAGATTAATTAATTAATTAAGACTGCCAGATCGAAGAACTACTTAAGATCAG 2923
QY      1218  -----LeuAlaValaGlnProMetAlaValaen----- 1227
Db      2924  AATTATATTGAAAATGCTCTCCGCTGCGACAGACAGTAGAGCCAAACTCAACGAAAGCT 2983
QY      1228  ---IleSerAlaenGlyPheGlyPheAlaIaIeCysGlnLeuSerValValTyrAnVa 1246
Db      2984  CACTGTGTCAACATCAGATGATGAGAAAAGCAACATGACCAATTTTGAACGATCTATAATGC 3043
QY      1246  lLyeAlaSerCylSerSerAArgAArgSerIleGlnAsnGlnGlnAlaPheAspLe 1266
Db      3044  ACAATTG-----AGGAGAGATGCAAATGTTTGCACAA-----TTCCACT 3085
QY      1266  uAspValAlaVal-----LysGluAsnLysAspAspLeuSerAnH1 1279
Db      3086  TGAATGTTTCTGTGAAGAAGCGACGATGAATTAAGAAGGCAAGGAGCCAAAGGAGC 3145
QY      1279  sValAspLeuSerValCysThrSerPheSerGlyProGlyArgSerGlyMetAlaLeu 1299
Db      3146  CCTCAAGCTCAAAATCTGCAGTACTGAGTATCTGGGGAAGATTGATTTCTCAATGACAAATAT 3205
QY      1299  tGluValAsnLeuSerGlyPheMetValProSerGluAlaIaIeSer---LeuSerG1 1318
Db      3206  TGAGTGTTCATMGCTGACTGGTATTGTCCCTGATATCTGAAGACCTTACGAGGCTTTCTAA 3265
QY      1318  uThrValLysLys-----ValGluTyrAspHis-----GlyLys 1329
Db      3266  AGAGAGTCACAGATATATCTCCATGTTTGAATTAACATATATATGGCTCAGAAAGGAC 3325
QY      1329  sLeuAsnLeuTyrLeuAspSerValaenGluThrGln---PheCysValaIaenIleProAl 1348
Db      3326  TGTATATCATTTACTTAAGCAAGAGTCTCCCACTCGAAGATGAATAGCTCGACCTTAAGAT 3385
QY      1348  aValArgAsnPheLysValaSerAnthrGlnAspAlaSerValaSerIleValaAspTyrTy 1368
Db      3386  TCTCAAGCATTTTGAAGTTGGCTTCATTCAGCCAGATCAGTCAAGGTGACAGCTACTA 3445
QY      1368  rGluProArgArgGlnAlaValaArgSerTyrAnSerGluValLysLeuSerSerCysAs 1388
Db      3446  CAATCTAGATGAAAAATGTACCAAGATCTACCATCCAGATGAGCAACAGGCTTCTCA 3505
QY      1388  pLeuCysSerAspValaGlnGlyCysArgProCysGluAspGlyAlaSer 1404
Db      3506  TAAAGTATATGTTGTGTAACGTTGCCGATGTCAGAAAGAAACGTTGCC 3554

RESULT 12
US-09-017-947-33
/ Sequence 33, Application US/09017947
/ Patent No. 6303754
/ GENERAL INFORMATION:
/ APPLICANT: VOGEL, CARL-WILHELM
/ APPLICANT: BREDEHORST, REINHORST
/ APPLICANT: KOCK, MICHAEL
/ APPLICANT: FRITZINGER, DAVID
/ TITLE OF INVENTION: RECOMBINANT PROCVF
/ NUMBER OF SEQUENCES: 39
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: OBION, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,

```



QY 416 -----Glnlyslleasn-----TyrThrValProGln----- 424  
 Db 16768 CGATGCCGCAATACGATACGCTTCGCTTATACCGACCTTCAGGTGTGAGGCC 16709  
 QY 425 -----SerGlyThrPheLysIleGluPheProIleLe 435  
 Db 16708 TGCCTCCGGCCGATGCGGCAAGTTCCTTCGGGACAGTACCGATGCGTTTGGCGAG 16649  
 QY 435 u-----GluAspSerSerGluLeuGlnLeuLysAlaTyrPheLeuGlySerLysSerSe 453  
 Db 16648 TGGCGGAGAGGATTCGCACTTGTTCAGCAAGACGTTATCTTCGCTCGCAAGATGCG 16589  
 QY 453 rmeAlaValHisSerLeuPheLysSerProSerLys-----ThrTyrIleGlnLe 470  
 Db 16588 TCAC-----TTGTGCAATCCCTCCCAAGGCTTTGCAATATGTGGCGGA 16544  
 QY 470 uLysThrArgAspGlnLysn-----IleLysValGlySerPro----- 482  
 Db 16543 GGAGAAATACAGACAGTACCGCCCGCCGATCCTGGTGGTCTACTCCCAAGATGCG 16484  
 QY 483 -----PheGluLeuValValSerGlyLysnLysValGlyLeu 494  
 Db 16483 GGATGCTCAGAGTGTATCTGTCTTACGACCTGACTCAGGAGG-----CGATTTCAT 16430  
 QY 494 sglLysSerTyrMetValValSerArgGlyGlnLeuValAlaValGlyLysGlnAnse 514  
 Db 16429 CGAAAGAGAGATGATCGCTTCGCGCGCGGAGATCGTCAAGATA-----GT 16382  
 QY 514 rThrMetPheSerLeuThrProGluAnSerTyrThrProLysAlaCyValIleValIly 534  
 Db 16381 GCCCATGCTCCGCGACGCTCG-----TTGCGCGAGATGATGATGTTTCTCT 16334  
 QY 534 rTyrIleGlnAspArgIleGluIleIleSerAspValLeuLysIleProValGlnLys 554  
 Db 16333 CTATTATGATATATGAGGAGGAGGCTGTATAGTCGAAGCTAGATCTGAAGCCCAAGATGCC 16274  
 QY 554 lPheLysAsnLysIleLysLeuTyrTyrPheSer-----LysValIlyAsnLysGluProSerG 572  
 Db 16273 C--AAGGAGAGATTAAGATGCTGTGGATGACTTTCGCGACAGATCCGACCTGGTGA 16217  
 QY 572 uLysValSerLeuArgIleSerValIleThrGlnProAspSerIleValGlyIleValAla 592  
 Db 16216 GAAGAGAGATGAGATGCTCTCTCCGCGATCGGAG-----GGCAA 16175  
 QY 592 lAspLysSerValAsnLeuMetAsnAla--SerAsnAspIleThrMetGlnAnVal 611  
 Db 16174 ACCGCTCTCATACATGATGAGCTGCTGATGATGATGATGCTCTGATATAGATCGT 16115  
 QY 611 lHis----- 612  
 Db 16114 GCACAGTCTGTTTCTCTCCGATTCCTCTTCTTTCGCGACTCCGCGCTCTTGTGT 16055  
 QY 613 -GluLeuGlnLeuLysThrAsnThrGlyTyrTyrLeuGlyMetPheMetLysSer 629  
 Db 16054 TAGGTTGCTTTCGCGATACGAGGATTCAGGTGCGGAGGTTTGTGAGAGATCTTACGT 15995  
 QY 630 -----PheAlaValPheGlnLysCysGlyLeuTyrValLe 641  
 Db 15994 GCAAGTCCGAGTTCAAACACCTGTTTCAAAACTCCCACTGCG--CTGTGCTGCG 15938  
 QY 641 uThrAspAlaAsn--LeuThrLysAspTyrIleAspGlyValTyrAsp 656  
 Db 15937 TACCGTTATAGATATGATGCTTCGAGATTAT-----TATATAGATGAGCCGATGAT 15887  
 QY 657 -----AsnAlaGluTyrAlaGluThrPhe--MetGlnLys 668  
 Db 15886 ATTGTGGGATATGAGATTCGCGAAGCAAAATCGGTGATGCTTCCGAGTGAAGAGAA 15827  
 QY 668 nGluGlyHisIleValAspIleHisAspPheSerLeuGlySer----- 682  
 Db 15826 AGCTGCTCGGCGAGATGCGTAATGAAAGCAAACTCCGATTCGAAAGATCCGGGATTTCA 15767  
 QY 683 -SerProHisValArgLysHisPheProGluThrTyrIleThr-----LeuAspThr 699

Db 15766 AGAATCTCCGTCGATACCAATTCCTGAGACTGATTCCTTGACCCGAGATTCGTAC 15707  
 QY 699 rAsnMetGlyTyrArgIleTyrGlnLysPheGluValIleThrValProAspSerIleThrSe 719  
 Db 15706 CGATGAACGGGGGAGAGTTTCTTGATCTTTC-----ACTCTCCGAGACATGACGCG 15653  
 QY 719 rThrValAlaThrGlyPheValIleSerGluAspLeuGlyLeuGlyLeuThrThrThr 739  
 Db 15652 CTGGACCTGCTCTCTTTCACATACAGAGATATGCGGTGGGATGGAACAGAAAG 15593  
 QY 739 oValGluLeuGlnAlaPheGlnProPhePheIlePheLeuAsnLeuProTyrSerValI 759  
 Db 15592 CGTGGAGCTGCG-----AAAGACTTATGCTTACACCAATTCGACGATCTCTGAG 15539  
 QY 759 eArgGlyGlnLysPheAlaLeuGluIleThrIlePheAsnTyrLeuLysAspAlaThrG 779  
 Db 15538 AATGGGTGACAAAGGTACCGCTTCGCTTCGATTCGGAACGGGAGGAGCGATGACGA 15479  
 QY 779 uValLysValIleIleGluLysSerAspLysPheAspIleLeuMetThrSerSerGlu 799  
 Db 15478 GGGCTTCGTCGATGGAAGCTTTCGAC-----CTGCTACCGCAA 15437  
 QY 799 eAsnAlaThrGlyHisGlnIleThrLeuLeuValProSerGluAspGlyAlaThrValLe 819  
 Db 15436 GCTATTGGGGGAGAGGCTTCGCTTCGCGAAGCCGCTGATACGGTTCGCTG 15377  
 QY 819 uPheProIleArgProThrHisLysLeuGlyGluIleProIleThrValIleThrLys 839  
 Db 15376 GTTTCGCTCGATCTCTGTC----- 15358  
 QY 839 oThrAlaSerAspAlaValThrGlnMetIleLeuValLysAlaGluGlyIleGluLysSe 859  
 Db 15357 -TCCGATATGATGCTCTGCTGCTGATGCTTGTATACCGAAAGCAAGCTTACCGATGG 15299  
 QY 859 rTyrSerGlnSerIleLeuLeuAspLeuThrAspAsnArgLeuGlnSerThrLeuLys 879  
 Db 15298 AGAGCAACATCTGATGCTGCACTGCGCGGCTACGAGCGAGTGGAGACCATACCTGT 15239  
 QY 879 rLeuSerPheSerPheProAsnThrValIleThrGlySerGluArgValGlnIleThrAl 899  
 Db 15238 GATTCTTCAGCGGACAGTCCCAACGGTG----- 15208  
 QY 899 aIleGlyAspValLeuGlyProSerIleAsnGlyLeuAlaSerLeuIleArgMetProTy 919  
 Db 15207 -----GATCTGAGATGCTA-----TTTCCCA 15185  
 QY 919 rGlyCysGlyGlnLysAsnMetIleAsnPheAlaProAsnIleTyrIleLeuAspTyrLe 939  
 Db 15184 TCGCAGCGGAGACCGGCTATGGGTACATGAGCACTTCAAGTG--GTAGCAATCCGCT 15128  
 QY 939 uThrLysLysLysGln-----LeuThrAspAsnLysLysGlyLysAlaLeuSe 955  
 Db 15127 TTGGGTGCGCGTACAGGCTTCGCGCGATGATGATGATGATGATGATGATGATGATGATG 15068  
 QY 955 rPheMetArgGlnGlyTyrGlnArgLysLeuLysTyrGlnArgLysLysSerPheSe 975  
 Db 15067 GGTGGCTTCGCGCTTATGCTCAATAGATATGCTGCTGCTGCTGCTGCTGCTGCTGCTG 15009  
 QY 975 rAlaPheGlyAsnTyr-----AspProSerGlySerThrTyrLeuSerAlaPheVal-- 992  
 Db 15008 TCGATCTCTGCTGCTGGTGAAGATT--TGGGACATCTTGGGCGGACCTTGGAGATCA 14950  
 QY 993 -----LeuArgCysPhe 996  
 Db 14949 CTCGCTTACAAATATGCGCTTGTCTGCGAGCAATGCGCTTGGCGCGGAGATGCTG 14890  
 QY 996 eLeuGlnAlaAspProTyrIleAspIleAspGlnAsnValLeuHisArgThrTyrThrTyr 1016  
 Db 14889 GCC-GAAGCAACATGCTGAGGCTTTCAGGCTTTCGCTGCTGCTGCTGCTGCTGCTGCTG 14831  
 QY 1016 r-----LeuLysGlyLysGlnLysSerAsnGlyLysPhe-- 1027



```

Db      14830 GACGAGCAGATGCTGACTGACAAACTGAAAGGCTGCAGAAAGTCTGACGGTTGGGGGC 14771
Oy      1028 -TTPASpProGlyArgValIleHisSerGluLeuGlnGlyIAsnLysSerProValIth 1047
Db      14770 TTGG-----CATCCTGAGATGTCCTTCAACGAT-----TA 14741
Oy      1047 rIeuThralaTyrlIeValIhrSerLeuLeu 1057
Db      14740 TCTGACGCACTAGTGATGACGAGTGTGTA 14710
RESULT 14
US-09-311-352B-1
; Sequence 1, Application US/09311352B
; Patent No. 6329500
; GENERAL INFORMATION:
; APPLICANT: Webb, Donna J.
; APPLICANT: Gonias, Steven L.
; TITLE OF INVENTION: Transforming Growth Factor-beta Binding Site
; FILE REFERENCE: 00370-02
; CURRENT APPLICATION NUMBER: US/09/311,352B
; CURRENT FILING DATE: 1999-05-13
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 339
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-311-352B-1

```

Alignment Scores:	
Pred. No.:	7,98e-16
Score:	234.00
Percent Similarity:	58.56%
Best Local Similarity:	45.95%
Query Match:	3.18%
DB:	4.18%
Length:	339
Matches:	14
Conservative:	15
Mismatches:	36
Indels:	3
Gaps:	3

US-10-020-095-4 (1-1428) X US-09-311-352B-1 (1-339)

Qy	660	TyAlGIuIarPhemeIcIuIaengIuGIyH:siIeValaIpIiH:saSPheSer	679
Db	28	TATGAGTCAGATGTATATGGGAAGAGCCATGCACCCCTGTGTCACTTGAAGAG-	81
Qy	680	LeuGIySerSerProH:s-----ValArGIySH:siPheProGIuThrTTrIleTrp	696
Db	82	-----CCTCACACGAGACCCGTACCAAAAGTACTTCCCTCGAACAACATGTGATCTGG	129
Qy	697	LeuAsPThrAsmMetGIyTrArGIleTrGIuGIuPheGIuValaThrValProAspSer	716
Db	130	GATTGTGTGTGTATAACTCAGCAGGGGTGCTGAGGTAGAGAGThAACGTCCCTCGAACCC	188
Qy	717	IleThSerTrpValAlaThrGIyPheValIleSerGIuAspLeuGIyLeuGIyLeuThr	736
Db	190	ATACCCGAGTGAAGGACGAGGGCTTGTGCCCTGTCAAGATGTCTGGACTTGATATCTCT	249
Qy	737	ThrThProValGIuLeuGIuAlaPheGIuProPhePheIlePheLeuAsmLeuProTr	756
Db	250	TTCACCT---GCCCTCTCCGAGGCTTCCAGCCCTCTTGTGTGAGCTACCAATGCCCTTAC	306
Qy	757	SetValIleArGIyGIuGIuPheAlaLeuGIu	767
Db	307	TCTGTATTCGTGTGAGAGGCTTCCACCTCAAG	339

RESULT 15  
 US-09-241-606-3  
 ; Sequence 3, Application US/09241606  
 ; Patent No. 6472140  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Tanzi, Rudolph E.  
 ; APPLICANT: Kovacs, Dora  
 ; APPLICANT: Saunders, Aleister J.  
 ; TITLE OF INVENTION: Alpha-2-Macroglobulin Therapies and Drug Screening Methods for  
 ; TITLE OF INVENTION: Alzheimer's Disease

```

: FILE REFERENCE: 0609.4460003
: CURRENT APPLICATION NUMBER: US/09/241,606
: CURRENT FILING DATE: 1999-02-02
: NUMBER OF SEQ ID NOS: 27
: SOFTWARE: PatentIn Ver. 2.0
: SEQ ID NO 3
:
: LENGTH: 750
:
: TYPE: DNA
:
: ORGANISM: Homo sapiens
:
: FEATURE:
:
: NAME/KEY: CDS
: LOCATION: (1)..(750)
:
: OTHER INFORMATION: A /LRP Binding Domain
US-09-241-606-3

```

Alignment Scores:	
Pred. No.:	3.08e-13
Score:	215.50
Percent Similarity:	45.08%
Best Local Similarity:	26.64%
Query Match:	2.93%
DB:	4
	750
Length:	750
Matches:	65
Conservative:	45
Mismatches:	95
Indels:	39
Gaps:	7

US-10-020-095-4 (1-1428) X US-09-241-606-3 (1-750)

Qy	1164	1166	1168	1170	1172	1174	1176	1178	1180	1182	1184	1186	1188	1190	1192	1194	1196	1198	1200	1202	1204	1206	1208	1210	1212	1214	1216	1218	1220	1222	1224	1226	1228	1230	1232	1234	1236	1238	1240	1242	1244	1246	1248	1250	1252	1254	1256	1258	1260	1262	1264	1266	1268	1270	1272	1274	1276	1278	1280	1282	1284	1286	1288	1290	1292	1294	1296	1298	1300	1302	1304	1306	1308	1310	1312	1314	1316	1318	1320	1322	1324	1326	1328	1330	1332	1334	1336	1338	1340	1342	1344	1346	1348	1350	1352	1354	1356	1358	1360	1362	1364	1366	1368	1370	1372	1374	1376	1378	1380	1382	1384	1386	1388	1390	1392	1394	1396	1398	1400	1402	1404	1406	1408	1410	1412	1414	1416	1418	1420	1422	1424	1426	1428	1430	1432	1434	1436	1438	1440	1442	1444	1446	1448	1450	1452	1454	1456	1458	1460	1462	1464	1466	1468	1470	1472	1474	1476	1478	1480	1482	1484	1486	1488	1490	1492	1494	1496	1498	1500	1502	1504	1506	1508	1510	1512	1514	1516	1518	1520	1522	1524	1526	1528	1530	1532	1534	1536	1538	1540	1542	1544	1546	1548	1550	1552	1554	1556	1558	1560	1562	1564	1566	1568	1570	1572	1574	1576	1578	1580	1582	1584	1586	1588	1590	1592	1594	1596	1598	1600	1602	1604	1606	1608	1610	1612	1614	1616	1618	1620	1622	1624	1626	1628	1630	1632	1634	1636	1638	1640	1642	1644	1646	1648	1650	1652	1654	1656	1658	1660	1662	1664	1666	1668	1670	1672	1674	1676	1678	1680	1682	1684	1686	1688	1690	1692	1694	1696	1698	1700	1702	1704	1706	1708	1710	1712	1714	1716	1718	1720	1722	1724	1726	1728	1730	1732	1734	1736	1738	1740	1742	1744	1746	1748	1750	1752	1754	1756	1758	1760	1762	1764	1766	1768	1770	1772	1774	1776	1778	1780	1782	1784	1786	1788	1790	1792	1794	1796	1798	1800	1802	1804	1806	1808	1810	1812	1814	1816	1818	1820	1822	1824	1826	1828	1830	1832	1834	1836	1838	1840	1842	1844	1846	1848	1850	1852	1854	1856	1858	1860	1862	1864	1866	1868	1870	1872	1874	1876	1878	1880	1882	1884	1886	1888	1890	1892	1894	1896	1898	1900	1902	1904	1906	1908	1910	1912	1914	1916	1918	1920	1922	1924	1926	1928	1930	1932	1934	1936	1938	1940	1942	1944	1946	1948	1950	1952	1954	1956	1958	1960	1962	1964	1966	1968	1970	1972	1974	1976	1978	1980	1982	1984	1986	1988	1990	1992	1994	1996	1998	2000	2002	2004	2006	2008	2010	2012	2014	2016	2018	2020	2022	2024	2026	2028	2030	2032	2034	2036	2038	2040	2042	2044	2046	2048	2050	2052	2054	2056	2058	2060	2062	2064	2066	2068	2
----	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	---

Qy	1377	SeTYrAeNSe	1380
Db	712	GAGTACATGCT	723

Search completed: January 17, 2004, 01:18:53  
Job time : 445 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: January 17, 2004, 01:19:00 ; Search time 982 Seconds  
(without alignments)  
5125.623 Million cell updates/sec

Title: US-10-020-095-4

Perfect score: 7348  
Sequence: 1 MOGPPLTAHLICVCTAAL.....HSSVIFPCFKLYMEIWL 1428

Scoring table: BLOSUM62  
Xgapop 10.0, Xgapext 0.5  
Ygapop 10.0, Ygapext 0.5  
Fgapop 6.0, Fgapext 7.0  
Delop 6.0, Delext 7.0

Searched: 2324096 seqs, 1762381658 residues

Total number of hits satisfying chosen parameters: 4648192

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlp  
-Q=/cgn2\_1/USPTO.epool\_p/US10020095/runatc\_15012004\_163824\_18739/app\_query.fasta\_1.1607  
-DB=PubMed Applications NA -QMT=fastcap -SUFFIX=rmpb -MINMATCH=0.1  
-LOOPEL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62  
-TRANS=human40.cdi -LIST=45 -LOCALIGN=200 -THR SCORE=apct -THR MAX=100  
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pro -NORM=ext -HEAPSIZE=500 -MINLEN=0  
-MAXLEN=2000000000 -USRR=US10020095 @CGN\_1.136 @runatc\_15012004\_163824\_18739  
-NCPUB=6 -ICPU=3 -NO MAP -LARGESQUERT -NEG SCORES=0 -WAIT -DSPELCK=100  
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5  
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications NA:\*

1: /cgn2\_6/ptodata/2/pubna/US07\_PUBCOMB.seq:\*  
2: /cgn2\_6/ptodata/2/pubna/PCT\_NEW\_PUB.seq:\*  
3: /cgn2\_6/ptodata/2/pubna/US06\_NEW\_PUB.seq:\*  
4: /cgn2\_6/ptodata/2/pubna/US06\_PUBCOMB.seq:\*  
5: /cgn2\_6/ptodata/2/pubna/US07\_NEW\_PUB.seq:\*  
6: /cgn2\_6/ptodata/2/pubna/PCTUS\_PUBCOMB.seq:\*  
7: /cgn2\_6/ptodata/2/pubna/US08\_NEW\_PUB.seq:\*  
8: /cgn2\_6/ptodata/2/pubna/US09\_PUBCOMB.seq:\*  
9: /cgn2\_6/ptodata/2/pubna/US09\_PUBCOMB.seq:\*  
10: /cgn2\_6/ptodata/2/pubna/US09C\_PUBCOMB.seq:\*  
11: /cgn2\_6/ptodata/2/pubna/US09C\_PUBCOMB.seq:\*  
12: /cgn2\_6/ptodata/2/pubna/US09\_NEW\_PUB.seq:\*  
13: /cgn2\_6/ptodata/2/pubna/US10A\_PUBCOMB.seq:\*  
14: /cgn2\_6/ptodata/2/pubna/US10A\_PUBCOMB.seq:\*  
15: /cgn2\_6/ptodata/2/pubna/US10B\_PUBCOMB.seq:\*  
16: /cgn2\_6/ptodata/2/pubna/US10\_NEW\_PUB.seq:\*  
17: /cgn2\_6/ptodata/2/pubna/US60\_NEW\_PUB.seq:\*  
18: /cgn2\_6/ptodata/2/pubna/US60\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
------------	-------	-------------	-----------	----	-------------

1	4535.5	61.7	3033	10	US-09-833-381-1810	Sequence 1810, App
2	3360	45.7	2273	12	US-10-108-260A-953	Sequence 953, App
3	1441	19.6	4576	13	US-10-292-081A-7	Sequence 7, Appli
4	1441	19.6	4577	14	US-10-052-817-1	Sequence 1, Appli
5	1440	19.6	4422	9	US-09-873-403-4	Sequence 4, Appli
6	1440	19.6	4577	9	US-09-873-403-3	Sequence 3, Appli
7	1440	19.6	4577	10	US-09-880-107-2236	Sequence 2236, Ap
8	1440	19.6	4577	12	US-10-331-496A-19	Sequence 19, Appl
9	1440	19.6	4577	13	US-09-873-319-08	Sequence 408, App
10	1440	19.6	4577	13	US-09-960-106-654	Sequence 654, App
11	1440	19.6	4577	13	US-10-282-081A-3	Sequence 3, Appli
12	1440	19.6	4577	13	US-10-292-081A-5	Sequence 5, Appli
13	1440	19.6	4577	15	US-10-076-816-56	Sequence 56, Appl
14	1440	19.6	4809	13	US-10-240-965-178	Sequence 178, App
15	1438	19.6	4527	13	US-09-756-247-3	Sequence 3, Appli
16	1438	19.6	4527	13	US-09-756-247-5	Sequence 5, Appli
17	1429	19.4	4530	13	US-10-282-081A-2	Sequence 2, Appli
18	1428	19.4	4488	13	US-09-981-151A-9	Sequence 9, Appli
19	1418	19.3	4823	13	US-09-971-392-172	Sequence 19, Appl
20	1412.5	19.2	4595	10	US-09-917-800A-1531	Sequence 1531, Ap
21	1410	19.2	5092	13	US-10-006-285-404	Sequence 404, App
22	1410	19.2	5092	13	US-09-971-429B-2	Sequence 2, Appli
23	1400	19.1	4615	10	US-09-880-107-3727	Sequence 3727, Ap
24	1390	18.9	4615	13	US-10-006-285-405	Sequence 405, App
25	1384	18.8	4677	13	US-10-316-253-266	Sequence 266, App
26	1384	18.8	4701	13	US-10-316-253-268	Sequence 268, App
27	1312	17.9	1300	13	US-10-133-013-223	Sequence 223, App
28	1274	17.3	4771	13	US-09-756-247-27	Sequence 27, Appl
29	1255	17.1	4421	9	US-09-925-301-552	Sequence 552, App
30	1157	15.7	4026	12	US-10-094-886-123	Sequence 123, App
31	1109.5	15.1	4501	12	US-10-094-886-125	Sequence 125, App
32	1072.5	14.6	5191	13	US-10-006-285-206	Sequence 206, App
33	991.5	13.5	5555	15	US-10-198-846-13979	Sequence 13979, A
34	988	13.4	5417	13	US-10-007-926A-413	Sequence 413, App
35	956	13.0	5087	15	US-10-001-076-111	Sequence 111, App
36	934	12.7	5067	11	US-09-964-824A-546	Sequence 546, App
37	934	12.7	5067	11	US-09-918-624B-18	Sequence 18, Appl
38	934	12.7	5067	13	US-10-101-510-316	Sequence 316, App
39	934	12.7	5067	15	US-10-001-076-3	Sequence 3, Appli
40	931	12.7	5067	9	US-09-875-519A-23	Sequence 23, Appl
41	930	12.7	875	13	US-10-160-162-48	Sequence 48, Appl
42	930	12.7	875	13	US-09-820-649-48	Sequence 48, Appl
43	930	12.7	5090	14	US-10-044-090-470	Sequence 470, App
44	925.5	12.6	6435	15	US-10-001-076-10	Sequence 10, Appl
45	921.5	12.5	4871	9	US-09-778-927A-18	Sequence 18, Appl

## ALIGNMENTS

RESULT 1  
US-09-833-381-1810  
; Sequence 1810, Application US/09833381  
; Patent No. US20020132090A1  
; GENERAL INFORMATION:  
; APPLICANT: Robison, Keith E.  
; TITLE OF INVENTION: No. US20020132090A1 Nucleic Acid and Protein Homologs  
; FILE REFERENCE: 5800-119  
; CURRENT APPLICATION NUMBER: US/09/833,381  
; PRIOR FILING DATE: 2001-04-11  
; PRIOR APPLICATION NUMBER: 09/516,448  
; PRIOR FILING DATE: 2000-02-29  
; NUMBER OF SEQ ID NOS: 2050  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 1810  
; LENGTH: 3033  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-833-381-1810

Alignment Scores:  
Pred. No.: 0 Length: 3033  
Score: 4535.50 Matches: 880  
Percent Similarity: 98.00% Conservative: 2

Best Local Similarity: 97.78% Mismatches: 1  
 Query Match: 61.72% Indels: 17  
 DB: 10 Gaps: 1  
 US-10-020-095-4 (1-1428) x US-09-833-381-1810 (1-3033)

QY 546 ValLeuYsIIeProValGlnLeuValPheLysAsnYsIIeYsLeuTYrTrpSerLys 565  
 Db 6 GTCCGAAAAATTCCTGTTACAGCTTCTTTTAAAAATAAGATTAAGCTATATTGGAGTAA 65  
 QY 566 ValLysAlaGluProSerGluLysValSerLeuArgIIeSerValThrnGlnProAspSer 585  
 Db 66 GTGAAAGCTGAACCATCTGAGAAAGCTCTCTTAGAGATCTGTGACACACCTGACCTCC 125  
 QY 586 IleValGlyIleValAlaValAspLysSerValAsnLeuMetAsnAlaSerAsnAspIle 605  
 Db 126 ATAGTTGGAGATTGTAAGCTGTGACAAAGGTGAAATCTGATGAATGCTCTTAAGATATT 185  
 QY 606 ThrMetGluAsnValAlaHISgluLeuGluLeuTYrAsnThrnGlyTYrTYrLeuGlyMet 625  
 Db 186 ACAATGAAAAATGTGTCATGATGAGTTGGAACCTTTATATACAGGATATTATTAGGCATG 245  
 QY 626 PheMetAsnSerPheAlaValPheGlnGluCysGlyLeuTYrValLeuThrAspAlaAsn 645  
 Db 246 TTCAATGAATCTTTTGGACGCTTTCCAGAAATGTGACTCTGGATATTGACAGATGCAAC 305  
 QY 646 LeuThrLysAspTYrIleAspGlyValTYrAspAsnAlaGluTYrAlaGluArgPheMet 665  
 Db 306 CTCACGAAAGATATTATGATGAGTGTATTATGACAAATGCAAAATGTGAGAGGATTATG 365  
 QY 666 GluGluAsnGluGlyHISIIeValAspIleHISAspPheSerLeuGlySerSerProHIS 685  
 Db 366 GAGGAAATGAAGGACATATTGTAGATATCATGACTTTCTTTGGGATGAGCATCCACAT 425  
 QY 686 ValArgLysHISpHeProGluThrTrpIleTrpLeuAspThrAsnMetGlyTYrArgIle 705  
 Db 426 GTCCGAAAGCATTTTCCAGAGACTGTGATTTGGCTGACACCAACATGGGGTTAACAGATT 485  
 QY 706 TyrGlnGluPheGluValThrValProAspSerIleThrSerTrpValAlaThrGlyPhe 725  
 Db 486 TACCAAGAAATTTGAAGTACTGACTGATCTGATCTGATCTGAGGAGCTACTGGTTTT 545  
 QY 726 ValIleSerGluAspLeuGlyLeuGlyLeuThrThrThrProValGluLeuGlnAlaPhe 745  
 Db 546 GTGATCTTGAGAGACCTGGCTTGGACTTAACAATCACTCCAGTGGAGCTCCAGGCTTC 605  
 QY 746 GlnProPhePheIlePheLeuAsnLeuProTYrSerValIleArgIlyGluGluPheAla 765  
 Db 606 CAACCATTTTTCATTTTGTGAAATCTTCCCTACTCTGTTATCAGAGGTGAAGAAATTTGCT 665  
 QY 766 LeuGluIleThrIlePheAsnTYrLeuLysAspAlaThrnGluValIleIleGlu 785  
 Db 666 TTGAAATAACTATATTAATTAATTTGAAAGATGCACTGAGGTTAAGTTAATCAATTGAG 725  
 QY 786 LysSerAspLysPheAspIleLeuMetThrSerSerGluIleAsnAlaThrnGlyHISGln 805  
 Db 726 AAAAGTGAACAATTTGATATCTAATGACTTCAATGAATAATTAATGCCAGGSCACAG 785  
 QY 806 GlnThrLeuLeuValProSerGluAspGlyAlaThrValLeuPheProIleArgProThr 825  
 Db 786 CAGAGCCCTTGTGTTCCAGTGAAGGATGGGGCAACTGTTCTTTTCCATCAGAGCCAAACA 845  
 QY 826 HIsleuGlyGluIleProIleThrValThrnAlaLeuSerProThrnAlaSerAspAlaVal 845  
 Db 846 CATGTGGAGAAATTCCTATACAGCTCACACTCTTTCACCCACTGCTTGAAGCTATC 905  
 QY 846 ThrGlnMetIleLeuValIleGluGlyIleGluLysSerTYrSerGlnSerIleLeu 865  
 Db 906 ACCAGATGATTTTAAAGCTGAAGGATGAAGAAATGAATAATCATTTCAATTCATCTTA 965  
 QY 866 LeuAspLeuThrAspAsnArgLeuGlnSerThrLeuTYrThrLeuSerPheSerPhePro 885  
 Db 966 TTAGACTTGACTGACAAATAGGCTACAGAGTACCTTGAAAACTTGAGTTTCTCATTTCT 1025

QY 886 ProAsnThrValThrnGlySerGluArgValGlnIleThrnAlaIleGluAspValLeuGly 905  
 Db 1026 CCTAATACACTGACTGGCAGATGAAGGTTCAATCACTACATCTCAATTTGGAGATGTTCTTGGT 1085  
 QY 906 ProSerIleAsnGlyLeuAlaSerLeuIleArgMetProTYrGlyCysGlyGluGlnAsn 925  
 Db 1086 CTTTCCATCAATGGCTTACGCTCATGATTCGATGCTTATGCTGCTGTGTAACGAAAC 1145  
 QY 926 MetIleAsnPheAlaProAsnIleTYrIleLeuAspTYrLeuThrLysLysGlnLeu 945  
 Db 1146 ATGATTAATTTCTCCAATTAATTTACATTTGATTAATGACTTAAGAAAGAAACAATG 1205  
 QY 946 ThrAspAsnLeuLysGluLysAlaLeuSerPheMetArgGlnGlyTYrGlnArgGluLeu 965  
 Db 1206 ACAGATTAATTTGAAGAAAGAAAGCTCTTTCATTTATGAGGCAAGGTTAACGAGAACTT 1265  
 QY 966 LeuTYrGlnArgGluAspGlySerPheSerAlaPheGlyAsnTYrAspProSerGlySer 985  
 Db 1266 CTCTATCAGAGGAAAGATGGCTCTTTCAGTGTGTTGGAAATATGACCTTCTGGGAGC 1325  
 QY 986 ThrTrpLeuSerAlaPheValIleuArgCysPheLeuGluAlaAspProTYrIleAspIle 1005  
 Db 1326 ACTTGTGTGACCTTTTGTTTAAGATGTTTCTTGAAACCGATCTTACATAGATATT 1385  
 QY 1006 AspGlnAsnValIleuHISArgThrTYrThrTrpLeuLysGlyHISgluLysSerAsnGly 1025  
 Db 1386 GATCAGAAATGTTTACACAGAACATACATCTGGCTTTAAAGAAATCAGAAATCCAAAGCT 1445  
 QY 1026 GluPheTrpAspProGlyArgValIleHISerGluLeuGlnGlyLysAsnLysSerPro 1045  
 Db 1446 GAATTTGGGATCCAGAAAGAGATTCATAGAGCTTCAAGGTGGCAATTAAGATCCCA 1505  
 QY 1046 ValThrLeuThrnAlaTYrIleValThrSerLeuLeuGlyTYrArgLysTYrGlnProAsn 1065  
 Db 1506 GTACACTTACACCTTATATGTAACCTCTCTGGAGATTAAGAAATGATCAGCCCTAAC 1565  
 QY 1066 IleAspValGlnGluSerIleHISpHeLeuGluSerGluPheSerArgGlyIleSerAsp 1085  
 Db 1566 ATTGATGTGAAAGCTATATCCATTTTGGAGTGTGAAATTCAGTGAAGGAATTTCAAGC 1625  
 QY 1086 AsnTYrThrLeuAlaLeuIleThrTYrAlaLeuSerSerValGlySerProLysAlaLys 1105  
 Db 1626 AATTATCTCTAGCCCTTATATCTTATGATGTGATCTGATGAGGAGCTCTTAAGCGAAG 1685  
 QY 1106 GluAlaLeuAsnMetLeuThrTrpArgAlaGluGlnGluGlyLysMetGlnPheTrpVal 1125  
 Db 1686 GAAGCTTGAATATGCTGACTTGGAGAGCAGAAAGAAAGTGCATGCAATTCGGGTG 1745  
 QY 1126 SerSerGluSerLysLeuSerAspSerTrpGlnProArgSerLeuAspIleGluValAla 1145  
 Db 1746 TCATCAGAGTCCAAACTTTTGACTCTCTGGACAGCAGCTCCCTGATATGGAAGTTGCA 1805  
 QY 1146 AlaTYrAlaLeuLeuSerHISpHeLeuGlnPheGlnThrSerGluGlyIleProIleMet 1165  
 Db 1806 GCGTATGCACTGCTTCACACTTCTTACAATTTCAAGCTTCAAGGGAATCCCAATTATG 1865  
 QY 1166 ArgTrpLeuSerArgGlnArgAsnSerLeuGlyGlyPheAlaSerThrnGlnAspThrTrn 1185  
 Db 1866 AGGTGTGTAAGCAGGAAAGAAATACCTGGGGGTGTTTGCACTCACTCAGATACCACT 1925  
 QY 1186 ValAlaLeuLysValaLeuSerGluPheAlaAlaLeuMetAsnThrnGlnArgThrnAsnIle 1205  
 Db 1926 GTGGCTTAAAGGCTGTGTGTAATTTGCAAGCCCTTAATGAATACAAAGAGCAATATATC 1985  
 QY 1206 GlnValThrValThrnGlyProSerSerProSerPro----- 1217  
 Db 1986 CAAGTACCGTGAAGGGGCTTACCTACCAAGTCTGTAAAGTTTCTGATTGACACACAC 2045  
 QY 1218 -----LeuAlaValAlaGlnProMetAlaValaAsnIle 1228  
 Db 2046 AACCGTTACTCTTCAGACAGACAGAGCTTGTGTGTACAGCAATGGCAGTAAATATT 2105

QY 1229 SerAlaAsnGlyPheGlyPheAlaIleCysGlnLeuAsnValValTyrAsnValIysAla 1248  
DB 2106 TCGCAAAAGTGTGGATTTGCTATTGTTCAGCTCAATGTTATATATGAAAGGCT 2165  
QY 1249 SerGlySerSerArgArgArgSerIleGlnAsnGlnGluAlaPheAspLeuAspVal 1268  
DB 2166 TCTGGATCTTTGAAAGAGAGAGATCTCAAAATCAAGAACCTTTGATTTAGATGTT 2225  
QY 1269 AlaValIysGlnAsnLysAspAspLeuAsnHisValAspLeuAsnValCysThrSerPhe 1288  
DB 2226 GCGTAAAGAAATAAATAGATGATCTCAATCAGTGGATTGATGTGTACAACTTT 2285  
QY 1289 SerGlyProGlyArgSerGlyMetAlaLeuMetGluValAsnLeuSerGlyPheMet 1308  
DB 2286 TCGGAGCCCGGTAGAGAGTGCATGAGCTTTATGAGATTAACTTAATAGTGGCTTTAG 2345  
QY 1309 ValProSerGluAlaIleSerLeuSerGluThrValIleValIysValGluTyrAspHisGly 1328  
DB 2346 GTGCTTCAGAAAGCAATTTCTGAGCAGAGACAGTAAAGTGAATATGATCATGGA 2405  
QY 1329 LysLeuAsnLeuTyrLeuAspSerValAsnGluThrGlnPheCysValAsnIleProAla 1348  
DB 2406 AAACCTCAACTCATTTAGATTTCTGTAATGAAACCACTTTTGTATTATTTCTGCT 2465  
QY 1349 ValArgAsnPhelyValSerAsnThrGlnAspAlaSerValSerIleValAspTyrTyr 1368  
DB 2466 GTGAGAACTTTAAAGTTTCAATATACCAAGATGCTTCAAGTCCATGAGATTACTAT 2525  
QY 1369 GluProArgArgGlnAlaValAlaArgSerTyrAsnSerGluValIleLeuSerSerCysAsp 1388  
DB 2526 GAGCCAAAGAGACAGCGGTGAGAAAGTTACACTGGAAGTGAAGCTGCTCCCTGTGAC 2585  
QY 1389 LeuCysSerAspValGlnGlyCysArgProCysGluAspGlyAlaSerGlySerHisHis 1408  
DB 2586 CTTTGAGAGATGTCAGGAGCTCCCTCTTGTGAGAGATGAGACTTCAGGCTCCATCAT 2645  
QY 1409 HisSerSerValIlePheIlePheCysPheValLeuLeuTyrPheMetGluLeuTyrLeu 1428  
DB 2646 CACTTTCAGTCACTTTTATTATTTCTGTTCAAGCTTCTGACTTTATGGAACCTTGGCTG 2705

RESULT 2  
US-10-108-260A-953  
; Sequence 953, Application US/10108260A  
; Publication No. US20040005560A1  
; GENERAL INFORMATION:  
; APPLICANT: HELIX RESEARCH INSTITUTE  
; TITLE OF INVENTION: No. US20040005560A1e1 full length cDNA  
; FILE REFERENCE: H1-A0106  
; CURRENT APPLICATION NUMBER: US/10/108,260A  
; NUMBER OF SEQ ID NOS: 5458  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 953  
; LENGTH: 2273  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-108-260A-953

Alignment Scores:  
Pred. No.: 0 Length: 2273  
Score: 3360.00 Matches: 657  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 45.73% Indels: 0  
DB: 12 Gaps: 0

US-10-020-095-4 (1-1428) x US-10-108-260A-953 (1-2273)

QY 1 MetGlnGlyProProleuLeuThrAlaAlaHisLeuLeuCysValCysThrAlaAlaLeu 20  
DB 66 ATGCAAGGCGCCACCGCTCCGACCGCGCCCACTTCTGCTGTGACCGCGCGCTG 125

QY 21 AlaValAlaProGlyProArgPheLeuValThrAlaProGlyIleIleArgProGlyGly 40

DB 126 GCCGTGGCTCCCGGAGCTTCGTGTTCTGGTGACAGCCCAAGGATCATCAGGCCCGAGAGA 185  
QY 41 AsnValThrIleGlyValGluLeuLeuGlnHisCysProSerGlnValThrValIysAla 60  
DB 186 AATGTGACTATTGGGAGTGGAGCTTCTGGAAACAGTCCCTTCAACAGTGACTGTGAAGCG 245  
QY 61 GluLeuLeuLysThrAlaSerAsnLeuThrValSerValLeuGluAlaGluGlyValPhe 80  
DB 246 GAGCTGCTCAAGACAGCATCAACCTCACTGTCTGTCTGTGAGAGCAAGAGAGATCTTT 305  
QY 81 GluLysGlySerPheLysThrLeuThrLeuProSerLeuProLeuAsnSerAlaAspGlu 100  
DB 306 GAAAAGGCTCTTTAAAGACACTTACTCTTCATCATCACTCTGAAACAGTGCAGATGAG 365  
QY 101 IleTyrGluLeuArgValThrGlyArgThrGlnAspGluIleLeuPheSerAsnSerThr 120  
DB 366 ATTTATGAGCTACGTTAAACCGGAGGTACCCAGATGAGATTTTATTTCTTAATGTATCC 425  
QY 121 ArgLeuSerPheGluThrTyrLeuArgIleSerValPheIleGlnThrAspLysAlaLeuTyr 140  
DB 426 CGCTTATCATTTGAGACCAAGAGATATCTGTCTTCAATCAACAGACAGAGCTTATAC 485  
QY 141 LysProLysGlnGluValIysPheArgIleValThrLeuPheSerAspPheLysProTyr 160  
DB 486 AAGCCAAAGCAAGAGTGAAGTTGCACTGTGTTACACTTCTCAGATTTTAAAGCTTAC 545  
QY 161 LysThrSerLeuAsnIleLeuIleLysAspProLysSerAsnLeuIleGlnGlnTyrLeu 180  
DB 546 AAACCTCTTTAAACATTTCTCAATTAAGAACCCCAATCAATTAATGATCAACAGGTGTTG 605  
QY 181 SerGlnGlnSerAspLeuGlyValIleSerLysThrPheGlnLeuSerSerHisProIle 200  
DB 606 TCACAAACAAAGTGAATCTTGAGACTCATTTCCAAACTTTTCACTTATCTTCCATTCATTA 665  
QY 201 LeuGlyAspTyrPheIleGlnValGlnValAsnAspGlnThrTyrTyrGlnSerPheGln 220  
DB 666 CTTGTGACGTGCTATTCAGTTCAAGTGAATGACCAAGACATCAATCAATCATTTGAG 725  
QY 221 ValSerGluTyrValLeuProLysPheGluValThrLeuGlnThrProLeuTyrCysSer 240  
DB 726 GTTTGAGATATGTATTAACAAATTTGAAGTGAATGACAGACACCATTAATATGTTCT 785  
QY 241 MetAsnSerLysHisLeuAsnGlyThrIleThrAlaLysTyrThrGlyLysProVal 260  
DB 786 ATGAATTTCAAGCATTTTAAATGATGATCAATCAGGCAAGATATCAATATGAGAGCCAGTG 845  
QY 261 LysGlyAspValThrLeuThrPheLeuProLeuSerPheTyrGlyLysLeysAsnIle 280  
DB 846 AAAGAGAGAGTAAAGCTTACATTTTACCTTATCTTTGGGAAAGAAAGAAATATTT 905  
QY 281 ThrLysThrPheLysIleAsnGlySerAlaAsnPheSerPheAsnAspGluLeuLys 300  
DB 906 ACAAACAACTTTAAGATTAATGATGATCTGCAAACTTCTTTTAATGATGAAGAGATGAAA 965  
QY 301 AsnValMetAspSerSerAsnGlyLeuSerGluTyrIleLeuAspLeuSerSerProGlyPro 320  
DB 966 AATGTAAATGATTTCTCAATGAGACTTTCGAAATACCTGAGATCAATCTTCCCTGGAACA 1025  
QY 321 ValGluIleLeuThrThrValThrGluSerValThrGlyIleSerArgAsnValSerThr 340  
DB 1026 GTAGAAATTTTAAACACAGTGAAGATCAAGTACAGGTATTTCAAGAAATGTAAGCACT 1085  
QY 341 AsnValPhePheLysGlnHisAspTyrIleIleGluPhePheAspTyrThrValLeu 360  
DB 1086 AATGTGTTCTTCAAGCAACATGATTAATCAATGATGATTTTGTATTAATCACTACTGTCCTG 1145  
QY 361 LysProSerLeuAsnPheThrAlaThrValIysValThrArgAlaAspGlyAsnGlnLeu 380  
DB 1146 AAGCATCTCTCAACTTACAGCCACTGTGAAAGGTAATCGTGAGTGGCAACCAACTG 1205  
QY 381 ThrLeuGluGluArgArgAsnAsnValValIleThrValThrGlnArgAsnTyrThrGlu 400

```

Db      1206 ACTCTTGAGAAAGAGAAATATGTACTCATTAACAGTGCACAGAGAACTATCTGAG 1265
Qy      401  TTTTTPSerGlySerAsnSerGlyAsnGlnlysmetGluAlaValGlnlyleAsnTyr 420
Db      1266 TACTCGACCGGATCTTAACAGTGCAGAAATGAGAAATGAGAGCTTTCAGAAATTAATTTAT 1325
Qy      421  ThrValProGlnSerGlyThrPheIysIleGlnPheProIleGlnGluAspSerSerGlu 440
Db      1326 ACTGCCCCCAAGTGAACCTTTTAAGATGAATCCCAATCTCGAGAGATTCCAGTGAAG 1385
Qy      441  LeuGlnLeuValAlaTyrPheLeuGlySerIysSerSerMetAlaValHisSerLeuPhe 460
Db      1386 CTACAGTTGAAGGCGCTATTTCTTGTGTAGTAAAGTACAGCAGGATTCATATGCTGTTT 1445
Qy      461  LysSerProSerIysThrTyrIleGlnLeuIysThrAspGluAsnIleLysValGly 480
Db      1446 AAGCTCTTCTAGTAAAGATACATCCAACTTAACAGAGATGAAATATTAAGGTGGGA 1505
Qy      481  SerProPheGlnLeuValValSerGlyAsnLysArgLeuLysGlnLeuSerTyrMetVal 500
Db      1506 TCGCTTTTGAAGTGTGTAGTGTAGTGCACAAAGATTGAAGAGATTAGCTATATGTA 1565
Qy      501  ValSerArgGlyGlnLeuValAlaValGlySerGlnAsnSerThrMetPheSerLeuThr 520
Db      1566 GTATCCAGGGGACAGTTGTGTGTGTAGGAAACAAATTCACAAATGTTCTTTTACA 1625
Qy      521  ProGluAsnSerTyrThrProLysAlaCysValIleValTyrTyrIleGluAspAspGly 540
Db      1626 CCGAAATTCCTTGACTCCAAAGCCCTGTGTATGTGTATATTAATTAAGATGATGGG 1685
Qy      541  GluIleIleSerAspValLeuLysIleProValGlnLeuValPheLysAsnLysIleLys 560
Db      1686 GAAATTAATTAAGTATGTTCTTAATAAATCTCTGTAGCTGTGTTTAAATAATTAAGTAAAG 1745
Qy      561  LeuTyrTyrSerLysValLysAlaGlnProSerGlyLysValSerLeuArgIleSerVal 580
Db      1746 CTATATTTGAGTAAAGTAAAGCTGAACCATCTGGAAGATCTCTTGAAGATCTCTGTG 1805
Qy      581  ThrGlnProAspSerIleValGlyIleValAlaValAspLysSerValAsnLeuMetAsn 600
Db      1806 ACACAGCTGACTCCATATGTTGGAGTGTGTAGCTGTGACAAAGTGCAGAACTGATGAT 1865
Qy      601  AlaSerAsnAspIleThrMetGluAsnValAlaHisGlnLeuGlnLeuTyrAsnThrGly 620
Db      1866 GCCCTATATGATTAATTAACATGGAAGAAATGTGTCATGATGTTGAACCTTAATTAACAGGA 1925
Qy      621  TyrTyrLeuGlyMetPheMetAsnSerPheAlaValPheGlnGluCysGlyLeuTyrVal 640
Db      1926 TATTATTATGAGCATGTCATGAATTTCTTTCAGTCTTTCAGGAATGTGAGCTCGGATA 1985
Qy      641  LeuThrAspAlaAsnLeuThrLysAspTyrIleAspGlyValTyrAspAsn 657
Db      1986 TTGACAGATGCAACCTTCACAGAGATTAATTAATGATGTGTTATGCAAT 2036

RESULT 3
US-10-292-081A-7
; Sequence 7, Application US/10292081A
; Publication No. US20030162202A1
; GENERAL INFORMATION:
; APPLICANT: Kenneth David Becker
; APPLICANT: Gonul Velicelbi
; APPLICANT: Xin Wang
; APPLICANT: Randolph E. Tanzi
; APPLICANT: Lars Bertam
; APPLICANT: Aleister J. Saunders
; TITLE OF INVENTION: SINGLE NUCLEOTIDE POLYMORPHISMS AND MUTATIONS ON ALPHA-2-MACROGLOBULIN
; FILE REFERENCE: 37481-3323
; CURRENT APPLICATION NUMBER: US/10/292,081A
; CURRENT FILING DATE: 2002-11-08
; PRIOR APPLICATION NUMBER: 60/337434
; PRIOR FILING DATE: 2001-11-09
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: FastSeq for Windows Version 4.0

```

```

; SEQ ID NO 7
; LENGTH: 4576
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-292-081A-7

Alignment Scores:
Pred. No.: 2,83e-133 Length: 4576
Score: 1441.00 Matches: 453
Percent Similarity: 46.20% Conservative: 264
Best Local Similarity: 29.19% Mismatches: 551
Query Match: 19.61% Indels: 284
DB: 13 Gaps: 60

US-10-020-095-4 (1-1428) x US-10-292-081A-7 (1-4576)
Qy      12  LeuLeuCysValCysThrAlaAlaLeuAlaValAlaProGlyProArgPheLeuValThr 31
Db      85  CTCTTGCTCTCTGCCCCACAGACGCGCTCAGTCTCTGGAAACCCGAGTATATGTTCTG 144
Qy      32  AlaProGlyIleIleArgProGlyGlyAsnValThrIleGlyValGlnLeuLeuHis 51
Db      145  GTCCCTCCCTGCTCCACACT---GAGACCACTGAGAGAGGAGTGTCTCTTCTGAGCTAC 201
Qy      52  CysProSerGlnValThrValLysAlaGlnLeuLeuIysThrAlaSerAsnLeuThrVal 71
Db      202  CTGAATGACACAGTACTGTATAGTGTCTCTTGGAGTGTCTGAGGGGAAACAGAGAGCTC 261
Qy      72  ---SerValLeuGlnAlaGlyLys---ValPheGlnLysGlySerPheLysThrLeuThr 89
Db      262  TTCACAGTACCTGAGGGGAGGAGATGACGTACTCCACTGTGTCCCTTC----- 309
Qy      90  LeuProSerLeuProLeuAsnSerAlaAspGlu-----IleTyrGlnLeuArgVal 106
Db      310  -----GCTGTCCCAAGTCTTATCAATCAAGAGAGTAAATGTTCTCACTGTCCAAAGT 363
Qy      107  ThrGlyArgThrGlnAspGlnLeuLeuPheSerAsnSerThrArgLeuSerPheGluThr 126
Db      364  AAGGACCAACCCAGAA-----TTTAAGAGGGGACCAACAGTATGTTAAGAAC 414
Qy      127  LysArgIleSerValPheIleGlnThrAspLysAlaLeuTyrLysProLysGlnGlnVal 146
Db      415  GAGAGACGTGTGTCTTGTGCCAGACAGCAACAACTACACAAACCAAGGGCAGACAGTG 474
Qy      147  LysPheArgIleValThrIlePheSerAspPheLysProTyrLysThrSerLeuAsnIle 166
Db      475  AAATTTGTGTGTCTCCATGATGAATACTTCAACCCCTGAAATGATGATTCACCTA 534
Qy      167  Leu---IleLysAspProLysSerAsnLeuIleGlnGlnThrPheLeuSerGlnSerAsp 185
Db      535  GTATACATTACAGATCCCAAGAAAGAAATCGCATGCGACAAATGGCAGATTCCAGTTAGAG 594
Qy      186  LeuGlyValIleSerLysThrPheGlnLeuSerSerHisProIleLeuGlyAspTyrSer 205
Db      595  GGTGGCTCAAGCAATTTCTTTTCCCTCTCAAGAGCCCTTCCAGGGGCTCTCAAG 654
Qy      206  IleGlnValGlnValAsnAsp-----GlnThrTyrTyrGlnSerPheGlnIleSerGlu 223
Db      655  GTGGTGTACAGAAATATCAGGTGGAAGGACAGACAC---CCTTCAACCGTGAAGAA 711
Qy      224  TyrValLeuProLysPheGlnValThrLeuGlnThrProLeuTyrCysSerMetAsnSer 243
Db      712  TTTGTTCTTCCCAAGTTTGAAGTACAAAGTAAACAGTGCACAAAGTATACACATCTTGAA 771
Qy      244  LysHisLeuGlnGlyThrIleThrAlaLysTyrThrTyrGlyLysProValLysGlyAsp 263
Db      772  GAAGAGATGAATATACAGTGTGTGCTTATACACATATGGAAGGCTGTGCTCGACAT 831
Qy      264  ValThrLeuThrPheLeuProLeu----- 271
Db      832  GTACCTGTAGCATTTGCAGAAAGTATAGTACGCTTCGACCTGCCAGGTTGAAGATTCA 891
Qy      272  -----SerPheTyrGlyLys 276

```



Db 892 CAGGCTTTCTGAGAAATTCAGTGCAGACAGCTAAACGCCAGGCTGCTTATTCAGCA 951  
 QY 277 LysIleAsnIleThrIleThyPheIleAsnIleSerIleAsnIlePheAsnIle 296  
 Db 952 GTAAAT-----ACCAAGCTTCTCAGCTGAAG-----AAGAG 987  
 QY 297 GluGluMetIleAsnValMetAsnSerSerAsnIleuSerIleuIleuPheLeuSer 316  
 Db 988 TATGAATGAAA-----CTTCACACTGAGGCCAGATCCAA 1023  
 QY 317 SerProGlyProValGluIleLeuThr-----ThrValThrIleuSerVal 331  
 Db 1024 GAAGAAAGCACTGTGTGAATGACTGGAAGGCACTGCCAGTAAATTCACAAACCAAA 1083  
 QY 332 ThrGlyIleSerArgAsnValSerThrAsnValPhePheIleGlnIleAspIleIle 351  
 Db 1084 ACCAAACTCTCA---TTGTGAAAGTGCACTCACTTCGACAGGGA-----ATT 1131  
 QY 352 GluPhePheAspIleThrValIleuIleProSerIleuAsnIleThrValIle 371  
 Db 1132 CCCTCTTTGGG-----CAAG 1146  
 QY 372 ValThrArgAlaAspGlyAsnGluLeuThrIleuGluGluArgArgAsnValValIle 391  
 Db 1147 GTGCGCTAGTATGATGGAAGGCGTCCCTATACCA-----ATAAAGTCATA 1194  
 QY 392 ThrValThrGluArgAsnIleThrGluIleuIleuSerGlySerAsnIleuIleu 411  
 Db 1195 TTCATC---AGAGAAATGAAAGCAACTATTAC---TCCATCTCCACCGATGAG 1245  
 QY 412 MetGluAlaValGln---LysIleAsnIleThr-----ValProGlnSerGlyThrPhe 428  
 Db 1246 CATGCGCTGTATAGATTCTCTATCAACACCAATGATGAGTACTCTTACTGTT 1305  
 QY 429 LysIleGluPhePro-----IleuGluAspSer 438  
 Db 1306 AGGCTCAATTCACAGAGATGATCCCTGTACGCTACAGAGGTGTGACAGAAACAC 1365  
 QY 439 SerGluLeuGluIleuLysAlaIlePheLeuGlySerLysSerMetAlaValIleSer 458  
 Db 1366 GAAGAGGCACATCACTGCTAT----- 1389  
 QY 459 LeuPheLysSerProSerIleuIleGlnLeuIleuThrArgAspGluAsnIleLys 478  
 Db 1390 CTGTGCTTCTCCCAAGCAAGAGCTTGTCCACTTGAGCCCATGTCTATGAATACCC 1449  
 QY 479 ValGlySer-----ProPheGluLeuValValSerGlyAsnLys---Arg 492  
 Db 1450 TGTGGCCATATCAAGACAGTCCAGCACAATTATTCGATGAGGACCCCTGCTGGG 1509  
 QY 493 LeuLysGluLeuSer-----TyrMetValValSerArgGlyGluLeuValAlaValGly 510  
 Db 1510 CTGAAGAGCTCTCTTATATATCGAATAATGCAAGGAGCAATTGTCCGAACTGGG 1569  
 QY 511 -----LysGluAsnSerThr-----MetPheSerLeuThr 520  
 Db 1570 ACTCATGACAGCTGTTGTGAAGCAGAGACATGAAGGCCATTTTCCATCTCAATCCCT 1629  
 QY 521 ProGluAsnSerIleThrProLysAlaCysValIleValTyrTyrIleGluAspArgIle 540  
 Db 1630 GTGAAGTACAGACTCTCTCGTCTGCTGCTGCTCATCATCATGCTTTTACTACCGG 1689  
 QY 541 GluIleIleSerAspValLeuLysIleProValGlnLeuValPheLysAsnLysIleLys 560  
 Db 1690 GAGTATTTGGGATTCGCAAAATATGATTTGAATAATGTCTGCGCAACAGGAGAT 1749  
 QY 561 LeuIleTyrSerIleValLysAlaGluProSerGluLysValSerLeuArgIleSerVal 580  
 Db 1750 TTGAGCTTCAGCCCATCAACAAATCTCCAGCTCAAGCCCACTGGAAGTCAACAGG 1809  
 QY 581 ThrGlnProAspSerIleValAlaValAlaValAspLysSerValAsnLeuMet--- 599

Db 1810 GCT---CTCAGTCCGCTGCGCCCTCCGCTGTGTGACCAAAAGGCTGCTCATGAAG 1866  
 QY 600 -----AsnIleSer----- 602  
 Db 1867 CCTGATGCTAGACTCTCGCGCTCTCGGTATCAACTCTGCTACCAAAAAGAACTCTACT 1926  
 QY 603 -----AsnAspIleThrMetGluAsnValValIleGluLeuLeu 616  
 Db 1927 GCGTTCCTCGGCGCTTTGATGATGACAGGACAAATGAAGACTGCATCATGCTATATGTC 1986  
 QY 617 TyrAsnThrGlyTyr---TyrIleuGlyMetPheMetAsnSer-----PheAla 631  
 Db 1987 TATATTAATGGAATCACAATATCTCCAGTATCACTACAAAGAAAGATATGTAACGC 2046  
 QY 632 ValPheGlnIleCysGlyLeuIleThrValLeuThrAspAlaAsnIleThrLysAspIle 651  
 Db 2047 TTCCTAGAGGACATGGGCTTAAAGGCAATTCACCAACTCAAGATTCGTAACCAAAATG 2106  
 QY 652 AspGlyValTyrAspAsnAlaGlu-----Tyr 660  
 Db 2107 TGTCCACAGCTTCAACAGATGAATAATGATGAGTACCTGAAGTCTACGTAGTTTAT 2166  
 QY 661 AlaGluArgPheMetGluGluAsnGluGlyIleValAspIleIleAspPheSerLeu 680  
 Db 2167 GAGTCAGATGTAAATGGAAGAGGCCATGCACGCTGTGTGCAATGTTGAAGAG----- 2217  
 QY 681 GlySerSerProIle-----ValArgLysIlePheProGluThrTyrIleuPhe 697  
 Db 2218 -----CCTCACACGAGACCTGACAAAGTACTTCCGACACATGATCTGGGAT 2268  
 QY 698 AspThrAsnMetGlyTyrArgIleTyrGlnIleuPheGluValThrValProAspSerIle 717  
 Db 2269 TTGGTGTGTGTAACTCAGACAGTGTGCTGAGTGAAGAGTACATGCTCCATACACATC 2328  
 QY 718 ThrSerTyrValAlaThrGlyPheValIleSerGluAspLeuGlyLeuGlyLeuThr 737  
 Db 2329 ACCGATGTGAAGGCGAGGCGCTTGTGCTGTGAAGATGTGAGCTTGATCTCTCC 2388  
 QY 738 ThrProValGluLeuGlnAlaPheGlnProPheIlePheLeuAsnLeuProIleThr 757  
 Db 2389 ACT---GCCTCTCTCGAGCTTCCAGGCTTCTTGTGTGAGCTCAACAATGCTTACTCT 2445  
 QY 758 ValIleArgGlyGluGluPheAlaLeuGluIleThrIlePheAsnIleLysAspAla 777  
 Db 2446 GGTATTCGTGAGAGGCGCTTCACTCAAGGCAAGGCTCTTAATCTCTCCCAAAATGC 2505  
 QY 778 ThrGluValLysValIleIleGluLysSerAspLysPheAspIleLeuMetThrSer 797  
 Db 2506 ATCCGGGTCAAGTGCAGCTGGAAGCTCTCCCGCTCTTCAAGTGTCCAGTGAAG 2565  
 QY 798 Glu-----IleAsnAlaThrGlyIleGlnIleThrLeuValProSer 812  
 Db 2566 GAACAAAGCGCTCACTGCATCTGTGCAACGCGCGGCA----- 2604  
 QY 813 GluAspGlyAlaThrValLeuPheProIleArgProThrIleLeuGlyGluIleProIle 832  
 Db 2605 -----ACTGTCTCTGCGCAGTAACCCCAAGTATTGGAATGTGAATTC 2652  
 QY 833 ThrValThrAlaLeuSer-----ProThrAla 841  
 Db 2653 ACTGTAGGCGGAGAGCACTAGAGTCTCAAGAGCTGTGTGGGACTGAGGCTTCACAGT 2712  
 QY 842 -----SerAspAlaValThrGlnMetIleLeuValAlaGluGlyIle 856  
 Db 2713 CCTGAACAGGAGGAAAGAAACACAGTCAACAGCTCTGTGTGTAACCTGAAGAGCTA 2772  
 QY 857 GluLysSerTyrSerGlnSerIleLeuLeuAspLeuThrAspAsnArgLeuGlnSerThr 876  
 Db 2773 GAGAAAGAAACAACATTCACTCTTGTCTCATAGTGTGAGAGTTC----- 2826  
 QY 877 LeuLysThrLeuSerPheSerPheProAsnThrValThrGlySerGluArgValGln 896  
 Db 2827 ---GAAGATATATCCCTGAACCTGCCCAATATGTGTGAAGATCTGCCGAGCTTCT 2883

QY 897 ILeThraLaIleGlyAspValIleuGlyProSerIleAsnGlyLeuAlaSerIleuIleArg 916  
 Db 2884 GTCCAGATTGGGAGCATATTAGGCTCTGCATGCAGAAACACAAACAAATCTTCTCAG 2943  
 QY 917 MetProGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGly 936  
 Db 2944 ATGGCTTATGGCTGGAGAGCAGAAATAGTCCCTTCTCTCAATATGATGATG 3003  
 QY 937 AspTyrIleuThrIleValGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGly 956  
 Db 3004 GATTATCTAAATGAACACAGAGCTTATCTCAGAGATCAAGTCCAGGCGCATGGCTAT 3063  
 QY 957 MetArgGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGly 976  
 Db 3064 CTCAAACAGCTGGTTACAGAGACAGTGAACAAACAGTATGCTGCTTACAGAGACC 3123  
 QY 977 PheGly---AsnTyrAspProSer---GlySerThrThrIleuSerAlaPheValIleuArg 994  
 Db 3124 TTTCGGGAGCGATATGGCAGAGAACCCAGGCAACCTGGCTCAGAGCTTGTCTGAG 3183  
 QY 995 CysPheIleuGlyAlaAspProTyrIleAspIleAsnValIleuIleAsnGlyThrTyr 1014  
 Db 3184 ACTTTGGCCAGCTCGAGCTCAATCTTCAATGATGAAGACACATTAACCAAGCCCTC 3243  
 QY 1015 ThrTyrIleuGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGly 1034  
 Db 3244 ATATGGCTCTCCAGAGCAGAGACAGACATGGCTGTTACAGAGCTCGGCTCACTGCTC 3303  
 QY 1035 HisSerGlyIleuGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGly 1054  
 Db 3304 AACATGCGCATTAAGGAGGAGTAGAAGATGAAGTCACTCTCGCTATATACCATC 3363  
 QY 1055 SerIleuGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGly 1074  
 Db 3364 GCCCTTCTGAGATTCTCTCAAGTCACTCACTGCTGCTGCTGCTGCTGCTGCTGCT 3423  
 QY 1075 LeuGlySer-----GluPheSerArgGlyIleSerAspAsnTyrThr 1088  
 Db 3424 CTGAGATGAGCTCGAAGACAGACAGACAGAGGAGCCATGGC---AGCATATATATAC 3480  
 QY 1089 LeuAlaIleuIleThrTyrAlaIleuSerSerValGlySer---ProValAlaValGlyAla 1107  
 Db 3481 AAACACAGCTGCGCTATGCTTTGCTCGAGAGTAAACAGCAAGAGAGAGAGAGTA 3540  
 QY 1108 LeuAsnMetIleuThrTyrAlaGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGly 1124  
 Db 3541 CTCAAGTCACTTAATGAAGAGCTGTGAAGAAACAACTGTCTCAATGGAGCGCCT 3600  
 QY 1125 ValSerSerGlySerIleuSerAspSerTyrGlnProArg-----SerLeuAspIle 1142  
 Db 3601 CAGAAACCAAGCAGCAGAGTGGGATTTTACAGAACCCAGGCTCCTGCTGAGAGTG 3660  
 QY 1143 GluValAlaAlaTyrAlaIleuSerSerIlePheIleuGlnPheGln-----ThrSer 1159  
 Db 3661 GAGATGACATCTTATGTCCTCTGCT---TATCTCAGGCGCCAGAGCCCAACCTCG 3717  
 QY 1160 GluGlyIlePro-----IleMetArgTyrIleuSerArgGlnArgAsnSerIleu 1175  
 Db 3718 GAGAGCTGAGCTCTGCAACCAACATCTGTGAAGTGAACAGAGCAGCAAAATGCCAG 3777  
 QY 1176 GlyGlyPheAlaSerThrGlnAspThrThrValAlaIleuValAlaIleuSerGluPheAla 1195  
 Db 3778 GCGCGTTTCTCTCCACAGAGACAGATGGTGGCTCTTCCATGCTGTCCAAATATGGA 3837  
 QY 1196 AlaLeuMetAsnThrGlyArgThrAsnIleGlnValThrValThrGlyProSerSerPro 1215  
 Db 3838 GCAGCACAATTATCC---AGGACTGGGAAGGCTGACAGCGATATCCATCTTCAAGG 3894  
 QY 1216 SerProLeuAlaValValGln-----1222  
 Db 3895 ACATTTTCCAGCAATTCAGATGAGACAAACAAACGCGCTGTATCTGACAGAGTCTCA 3954

QY 1223 -----ProMetAlaValAsnIleSerAlaAsnGlyPheGlyPheAlaIleCys 1238  
 Db 3955 TTGCGAGAGCTGCGGAGATACAGCATGAAGTAGACAGAGAGATGTGTACTCTC 4014  
 QY 1239 GlnLeuAsnValTyrAsnValIleValAsnGlySerSerArgArgAlaGlySerIle 1258  
 Db 4015 CAGACATCTCTGAATTAACATAT-----CTC 4041  
 QY 1259 GlnAsnGlnIleAlaPheAspLeuAspValAlaValIle-----GluAsnIleAsp 1275  
 Db 4042 CCAGAAAGAGAAAGTTCCTCCCTTGTAGAGTCAAGCTGCTGCTCAAACTGTGAT 4101  
 QY 1276 AspLeuAsn---HisValAspLeuAsnValCysThrSerPheSerGlyProGly----- 1292  
 Db 4102 GAACCCAAAGCCCAACAGCTTCAAAATCTCCCTAAGTGTCAATTAACAGAGAGCCGC 4161  
 QY 1293 ---ArgSerGlyMetAlaLeuMetGluValAsnIleuSerGlyPheMetValProSer 1311  
 Db 4162 TCTGCTCCAAATGGGAGATGTTGATGTAAGATGCTCTGCTTCAATTCCTGAG 4221  
 QY 1312 GluAlaIleSerIleu-----SerGluThrValIleValValGlyValAspPheIleGly 1328  
 Db 4222 CCAGACAGTAAATGCTTGAAGATTAACCATGATGAGCCGAGACAGAGACAGAGAC 4281  
 QY 1329 IysLeuAsnIleuTyrIleuAspSerValAsnGluThrGlnPheCysValAsnIleProAla 1348  
 Db 4282 CATGCTTGAATTAACCTTGTATGAGTGTCAATCAGACAGTGAAGTCTTCAAGCTT 4341  
 QY 1349 ValArgAsnPheIleValSerAsnThrGlnAspAlaSerValSerIleValAspTyrTyr 1368  
 Db 4342 CTCAGAGATGCCAGTGAAGATCTGAACCGACGATGTAAGTCTATGATTAATCTAC 4401  
 QY 1369 GluProArgGlnAlaValAlaArgSerTyrAsnSer 1380  
 Db 4402 GAGACGATGATGTTGCAATTCCTGAGTACAAATGCT 4437

RESULT 4  
 US-10-052-817-1  
 : Sequence 1, Application US/10052817  
 : Publication No. US20020114792A1  
 : GENERAL INFORMATION:  
 : APPLICANT: Tanzi, Rudolph E.  
 : APPLICANT: Kovacs, Dora  
 : APPLICANT: Saunders, Aleister J.  
 : TITLE OR INVENTION: Alpha-2-Macroglobulin Therapies and Drug Screening Methods for  
 : FILE REFERENCE: 0609,4460005  
 : CURRENT APPLICATION NUMBER: US/10/052,817  
 : PRIOR FILING DATE: 2002-01-23  
 : PRIOR APPLICATION NUMBER: 09/241,606  
 : PRIOR FILING DATE: 1999-02-02  
 : PRIOR APPLICATION NUMBER: 09/148,503  
 : PRIOR FILING DATE: 1998-09-04  
 : PRIOR APPLICATION NUMBER: 60/093,297  
 : NUMBER OF SEQ ID NOS: 27  
 : SOFTWARE: PatentIn Ver. 2.0  
 : SEQ ID NO 1  
 : LENGTH: 4577  
 : TYPE: DNA  
 : ORGANISM: Homo sapiens  
 : FEATURE:  
 : NAME/KEY: sig\_peptide  
 : LOCATION: (44)..(112)  
 : NAME/KEY: CDS  
 : LOCATION: (44)..(4468)  
 : NAME/KEY: mat\_peptide  
 : LOCATION: (113)..(4468)  
 : US-10-052-817-1

Alignment Scores: 2,83e-133 Length: 4577  
 Pred. No.: 1441.00 Matches: 451  
 Score:

Percent Similarity: 46.20% Conservative: 266  
 Best Local Similarity: 29.06% Mismatches: 551  
 Query Match: 19.61% Indels: 284  
 DB: 14 Gaps: 59

US-10-020-095-4 (1-1428) x US-10-052-817-1 (1-4577)

QY 12 LeuLeuCyValAlaLeuAlaValAlaProGlyProArgPheLeuValThr 31  
 |||||  
 |||||  
 DB 86 CTCCTGCTCTCTGCGCCACAGACGCTCAGTCTTGAAACCGAGATATATGTTCTG 145

QY 32 AlaProGlyIleLeuArgProGlyGlyAsnValThrIleGlyValGluLeuGluHis 51  
 |||||  
 |||||  
 DB 146 GTCCCTCCCTGCTCCACACT---GAGACCACTGAGAGAGGCTGCTCTTCGACCTAC 202

QY 52 CyProSerGlnValThrValIleValIleLeuLeuLeuThrAlaSerAsnLeuThrVal 71  
 |||||  
 |||||  
 DB 203 CTGATAGACACAGTGTACTGATGCTCTTCTGAGTCTGTACAGGAGAAACAGAGCCCTC 262

QY 72 ---SerValIleGluIleGluGly---ValPheGluIleGlySerPheLeuThr 89  
 |||||  
 |||||  
 DB 263 TTCACCTGACCTGGAGCGGAGATGACGTATCCACTGTGTGCGCTTC----- 310

QY 90 LeuProSerLeuProLeuAsnSerAlaSerGlu-----IleTyrGluLeuArgVal 106  
 |||||  
 |||||  
 DB 311 -----GCTGTCCCAAGTCTTCATCAATGAGAGATATGTTCTCTCACTGCTCAAGTG 364

QY 107 ThrGlyArgThrGlnAspGluIleLeuPheSerAsnSerThrArgLeuSerPheGluThr 126  
 |||||  
 |||||  
 DB 365 AAAGGCCAACCCCAAGAA-----TTTAAAGAACGGACCACTGATGTGTTAAGAC 415

QY 127 LysArgIleSerValPheIleGlnThrAspIleValAlaLeuTyrIlePheProIleGluVal 146  
 |||||  
 |||||  
 DB 416 GAGGACAGTGTGCTCTTGTCCAGACAGCAATCAATCACTAACACAGGACAGACAGT 475

QY 147 LysPheArgIleValThrLeuPheSerAspPheIlePheTyrIleThrSerLeuAsnIle 166  
 |||||  
 |||||  
 DB 476 AATATTGCTGTGTCTCATGTATGAATCACTTCACTCCCTGATGATGATTCCACTA 535

QY 167 Leu---IleLeuAspProIleSerAsnLeuIleGlnGlnThrLeuSerGlnIleAsp 185  
 |||||  
 |||||  
 DB 536 GTATACATTCAGAGATCCCAAGAGAAATCCGATCGCACAATGCGACAGTTTCCAGTTAAG 595

QY 186 LeuGlyValIleSerIleThrPheGlnLeuSerSerIleProIleLeuGlyAspTyrSer 205  
 |||||  
 |||||  
 DB 596 GGTGGCTCAAGCAATTTCTTTTCCCTCTCATTCAGAGCCCTTCAGAGGCTCTCAAG 655

QY 206 IleGlnValGlnValAsnAsp-----GlnThrTyrTyrGlnSerPheGlnValSerGlu 223  
 |||||  
 |||||  
 DB 656 GTGGTGTACAGAAATCAGGTGAGAGACAGACAC---CCTTTCACCGTGAAGAA 712

QY 224 TyrValIleLeuProIlePheGluValThrLeuGlnThrProLeuTyrCysSerMetAsnSer 243  
 |||||  
 |||||  
 DB 713 TTGTGTTCTTCCCAAGTTTGAAGTACAAAGTACAGTCCCAAGATATATCACTTTGAA 772

QY 244 LysHisLeuAsnGlyThrIleThrAlaLysTyrThrTyrGlyLysProValIleGlyAsp 263  
 |||||  
 |||||  
 DB 773 GAAAGATATATGTATCACTGTGTGCTTATACATATGGAGAGCTGTCCCTGACAT 832

QY 264 ValIleLeuThrPheLeuProLeu----- 271  
 |||||  
 |||||  
 DB 833 GTGACGTGAGCATTTGAGAAAGTATAGTACGCTTCGACGCGCAGGTAAGATTCA 892

QY 272 -----SerPheTyrGlyLys 276  
 |||||  
 |||||  
 DB 893 CAGGCTTTCTGTGAGAAATTCAGTGCAGACTTAACAGCCATGCTCTTATACAGCA 952

QY 277 LysLysAsnIleThrIleThrPheLysIleAsnGlySerAlaAsnPheSerPheAsnAsp 296  
 |||||  
 |||||  
 DB 953 GTRAAA-----ACCAAGGCTTTCAGCTGAAGAG-----AAGGAG 988

QY 297 GlnGluMetLysAsnValMetAspSerSerAsnGlyLeuSerGluTyrLeuAspLeuSer 316  
 |||||  
 |||||

DB 989 TATGAATGAAA-----CTTCACTAGAGCCCGATGCCAA 1024

QY 317 SerProGlyProValGluIleLeuThr-----ThrValThrGluSerVal 331  
 |||||  
 |||||

DB 1025 GAAAGAGAACAGTGTGATTTGACTGAAAGGCTCCAGTGAATATCAAGAAACATA 1084

QY 332 ThrGlyIleSerArgAsnValSerThrAsnValPhePheGlnHisAspTyrIleIle 351  
 |||||  
 |||||  
 DB 1085 ACCAAACTCTCA---TTTGTGAAGATGACTACACTTTCGACAGAG-----ATT 1132

QY 352 GluPhePheAspTyrThrThrValLeuLysProSerLeuAsnPheThrAlaThrValLys 371  
 |||||  
 |||||  
 DB 1133 CCTTCTTTGGG-----CAG 1147

QY 372 ValThrArgAlaAspGlyAsnGlnLeuThrLeuGluGluIleArgArgAsnValValIle 391  
 |||||  
 |||||  
 DB 1148 GTGCCCTTAGTATGATGAGAAAGGCTCCCTATAC-----ATAAATCATTA 1195

QY 392 ThrValThrGlnArgAsnTyrThrGluTyrTyrPheSerIleAsnSerGlyAsnGlnLys 411  
 |||||  
 |||||  
 DB 1196 TTCATC---AAGAGAAATGAAAGCAAACTATATAC-----TCAAATGCTAACCGGATGAG 1246

QY 412 MetGluAlaValGln---LysIleAsnTyrThr-----ValProGlnSerGlyThrPhe 428  
 |||||  
 |||||  
 DB 1247 CATGCCCTTGTATCACTTCTATCAACACCAACGTTATGGGTACCTCTTACTGTT 1306

QY 429 LysIleGluPhePro-----IleLeuGluAspSer 438  
 |||||  
 |||||  
 DB 1307 AAGGTCAATTACAAAGATGATGATCCCTGTTACGGCTTACAGTGGGTGTCAAGAAACAC 1366

QY 439 SerGluLeuGlnLeuLysAlaIleTyrPheLeuGlySerLysSerSerMetAlaValHisSer 458  
 |||||  
 |||||  
 DB 1367 GAAGAGGACATCAACACGCTTAT----- 1390

QY 459 LeuPheLysSerProSerLysThrTyrIleGlnLeuLysThrArgAspGluAsnIleLys 478  
 |||||  
 |||||  
 DB 1391 CTGTGTTCTCCCAAGAGAGCTTGTCCACCTTGAGCCCATGCTCATGATCACTACC 1450

QY 479 ValGlySer-----ProPheGluLeuValIleSerGlyAsnLys-----Arg 492  
 |||||  
 |||||  
 DB 1451 TGTGGCATACTACAGACAGTCCAGGACATTAATATTCGAATGAGAGGACCCCTGTGGGG 1510

QY 493 LeuLysGluLeuSer-----TyrMetValIleSerArgIleGlnLeuValAlaValGly 510  
 |||||  
 |||||  
 DB 1511 CTGAAGAGCTCTCTTTATATATCTGATATGCAAGAGAGGATTTGCGAAGTGGG 1570

QY 511 -----LysGlnAsnSerThr-----MetPheSerLeuThr 520  
 |||||  
 |||||  
 DB 1571 ACTCATGACCTGCTGTGAGAGAGAGACATGAAGGCCATTTTTCATCTCAATCCCT 1630

QY 521 ProGluAsnSerTyrThrProLysAlaCysValIleValIleTyrIleGluAspAspGly 540  
 |||||  
 |||||  
 DB 1631 GTGAAGTACAGCAATTCCTCTGTGCTGGTGTGCTCATCTATGCTGTTTACCTACCGGG 1690

QY 541 GluIleIleSerAspValLeuLysIleProValGlnLeuValPheLysAsnLysIleLys 560  
 |||||  
 |||||  
 DB 1691 GACGTGATTTGGGAGATTCGCAAAATATATGATTTGAAAATTTGTCGGCCAAACAGGTGAT 1750

QY 561 LeuTyrTyrPheSerLysValIleValIleGluProSerGluLysValSerLeuArgIleSerVal 580  
 |||||  
 |||||  
 DB 1751 TTGACCTTCAAGCCATCAACAAGTCTCCAGCTCACAGGCCACCTCGAGTCAACAGG 1810

QY 581 ThrGlnProAspSerIleValIleValIleValAlaIleAspLysSerValAsnLeuMetAsn 600  
 |||||  
 |||||  
 DB 1811 GCT---CCTCACTGCTGTGGGCTCTGCTGTGGAGCCAAAGCTGCTGCTCATGTAGAG 1867

QY 601 AlaSerAsnAspIleThrMetGluAsnValValIleGluLeu-----GluLeuTyrAsnThr 619  
 |||||  
 |||||  
 DB 1868 CCTGATGCTAGAGCTCTCGGCTCTCGGTTTACAACTGTACACGAAAGAGACCTCACT 1927

QY 620 GlyTyr-----TyrLeuGlyMet 625  
 |||||  
 |||||  
 DB 1928 GGCTTCCCTGGGCTTTGATATGACAGGACGATGAAGACTGCATCATGTATATGTC 1987

Qy 626 PheMetanSer-----PheAla 631  
 Db 1988 TATATTATATGGAATCATATATCTCACTATCATAGTACAATGAAAAGAGATATATGACGC 2047  
 Qy 632 ValPheGlnGluCyseGlyLeuThrValLeuThrAspAlaAsnLeuThrLysAspTyrIle 651  
 Db 2048 TTCTAGAGGACATGGCTTTAAAGGCATTACCACTCAAGATTCTGTTAAACCCAAATAT 2107  
 Qy 652 AspGlyValTyrAspAsnAlaGlu-----Tyr 660  
 Db 2108 TGTCCACAGCTTCAACAGTATGAAATGCAATGACCTGACCTTACGTTAGCTTTTAT 2167  
 Qy 661 AlaGluArgPheMetGluGluAsnGluGlyHisIleValAspIleHisAspPheSerLeu 680  
 Db 2168 GAGTCAGATGTTATGGAAGAGGCGCATGACGCCCTGGTCATGTTGAAAGAG----- 2218  
 Qy 681 GlySerSerProHis-----ValArgLysHisPheProGluThrTyrIleTyrLeu 697  
 Db 2219 -----CCTCACACGAGACCGCTACGAAAGTACTTCCCTGAGACATGATCTGGGAT 2269  
 Qy 698 AspThrAsnMetGlyTyrArgIleTyrGlnGluPheGluValThrValProAspSerIle 717  
 Db 2270 TTGGTGTGTAAACTCAGACAGGGGTGCTGAGGTAGAGTAACTCTTGACACCATC 2329  
 Qy 718 ThrSerTyrValAlaThrGlyPheValIleSerGluAspLeuGlyLeuGlyLeuThrThr 737  
 Db 2330 ACCAGATGGAAGGACGGGGCTTGTGCTGTGTAAGATGCTGACCTTGATATCTTCC 2389  
 Qy 738 ThrProValGluLeuGluAlaPheGlnProPhePheIlePheLeuAsnLeuProTyrSer 757  
 Db 2390 ACT---GCCTCTCCGAGCCCTTCCAGCCCTTCTTGAGAGCTTACCAATCCTTACTCT 2446  
 Qy 758 ValIleArgGlyGluGluPheAlaLeuGluIleThrIlePheAsnTyrLeuLysAspAla 777  
 Db 2447 GTGATTCGTGAGAGCCCTTACACTCAAGGCCAGGCTCTTAAAGTACTTCCCAATATGC 2506  
 Qy 778 ThrGluValLysValIleIleGlyLysSerAspLysPheAspIleLeuMetThrSerSer 797  
 Db 2507 ATCCGGGTCACTGTGCTCACTGGAACCTCTCCGCCCTTCTGCTGCTCCAGTGAAGAG 2566  
 Qy 798 Glu-----IleAsnAlaThrGlyHisGlnGlnThrLeuLeuValProSer 812  
 Db 2567 GAAACAAGCGCTCACTGATCTGTGCAAAACGGCGGCA----- 2605  
 Qy 813 GluAspGlyAlaThrValLeuPheProIleArgProThrHisIleGluGlyIleProIle 832  
 Db 2606 -----ACTGTCTCTGGGAGGTAAACCCAAAGTCAATTAGGAATGTGAATTTC 2653  
 Qy 833 ThrValThrAlaLeuSer-----ProThrAla 841  
 Db 2654 ACTGTGACCGCAGAGGACACTAGACTCAAGAGCTGTGGAGACGAGAGTCCCTTCACTT 2713  
 Qy 842 -----SerAspAlaValThrGlnMetIleLeuValLysAlaGluGlyIle 856  
 Db 2714 CCTGAACAAGAGAAAGACACACAGTCACTGTTGTTGTTGAACCTGGAAGAGACTA 2773  
 Qy 857 GluLysSerTyrSerGlnSerIleLeuLeuAspLeuThrAspAsnArgLeuGlnSerThr 876  
 Db 2774 GAGAAAGAAACACATTCACCTCTACTTGTTCATCAGGTGTGAGTGTCTT----- 2827  
 Qy 877 LeuLysThrLeuSerPheSerPheProAsnThrValThrGlySerGluArgValGln 896  
 Db 2828 ---GAAATTTATCTCTGAAACTGCAACCAATGTGTAAGAAATCTGCCGAGCTTCT 2884  
 Qy 897 IleThrAlaIleGlyAspValLeuGlyProSerIleAsnGlyLeuAlaSerLeuIleArg 916  
 Db 2885 GTCTCAGATTGGGAGACATATTAGCTCTGCAATGCAAAACACAAATCTTCTCAG 2944  
 Qy 917 MetProTyrGlyCyseGlyGluGlnAsnMetIleAsnPheAlaProAsnIleTyrIleLeu 936  
 Db 2945 ATGCCCTATGCTGTGGAGACAGAAATATGTCCTTGTCTCTTAAATCATCTATGACTG 3004

Qy 937 AspTyrLeuThrLysLysGlnLeuThrAspAsnLeuLysGluLysAlaLeuSerPhe 956  
 Db 3005 GATATCTTAATGAACAACAGACGCTTACTCCAGAGATCAAGTCCAAAGGCCATATGCTAT 3064  
 Qy 957 MetArgGlnIleTyrGlnArgGluLeuLeuTyrGlnArgGluAspGlySerPheSerAla 976  
 Db 3065 CTCAACACTGTTTACAGAGACAGTTGAATCTACAAACATATATATGCTCTTACAGCAC 3124  
 Qy 977 PheGly---AsnTyrAspProSer---GlySerThrTyrLeuSerAlaPheValLeuArg 994  
 Db 3125 TTGGGAGGAGATATGCGAGAACAGGACGACACCTGCTCACAGCTTGTTCGTAG 3184  
 Qy 995 CysPheLeuGluAlaAspProTyrIleAspIleAspGlnAsnValLeuHisArgThrTyr 1014  
 Db 3185 ACTTTTCCCAACTCGACGCTTACATCTTCATGATTAACACACATATTCCCAAGCCCTC 3244  
 Qy 1015 ThrTyrLeuLysGlyHisGlnLysSerAsnGlyGluPheThrAspProGlyArgValIle 1034  
 Db 3245 ATATGGCTCTCCAGAGCAGAAAGACATATGCTGTTTCAGAGCTTGAGTCACTGCTC 3304  
 Qy 1035 HisSerGluLeuGlnGlyAsnLysSerProValThrLeuThrAlaTyrIleValThr 1054  
 Db 3305 AACATGCCATTAAGGAGAGATGAAGATGAAGTACCTTCCGCTTATATCACATC 3364  
 Qy 1055 SerLeuLeuGlyTyrArgLysTyrGlnProAsnIleAspValGlnGluSerIleHisPhe 1074  
 Db 3365 GCCCTTGTGAGATTCCTCTCACAGTCACTACCCCTTGTCCGAATGCCCTGTTTGC 3424  
 Qy 1075 LeuGluSer-----GluPheSerArgGlyIleSerAspAsnTyrThr 1088  
 Db 3425 CTGAGTACACCTGGAAGACAGCAAGAAAGGAGCAATGCG---ACCATGTATATATAC 3481  
 Qy 1089 LeuAlaLeuIleThrTyrAlaLeuSerSerValGlySer---ProLysAlaLysGluAla 1107  
 Db 3482 AAAGCACTGCTGGCTTATGCTTTTGTGCTGCGAGTACACAGACAAAGAGAAAGTA 3541  
 Qy 1108 LeuAsnMetLeuThrTyrArgAlaGlnGluGlyGlyMetGlnPheTyr----- 1124  
 Db 3542 CTCAAGTCACTTATGAGAAAGCTGTGAAGAAAGACAACTGTGATGGAGCGCCT 3601  
 Qy 1125 ValSerSerGluSerLysLeuSerAspSerTyrGlnProArg-----SerLeuAspIle 1142  
 Db 3602 CAGAAACCAAGGACAGCAGTGGGCAATTTTACGAACCCAGGCTCCTGTGTGAGGTG 3661  
 Qy 1143 GluValAlaAlaTyrAlaLeuSerSerHisPheLeuGlnPheGln-----ThrSer 1159  
 Db 3662 GAGATGACATCTATGTGCTCTGCTGCT---TATCTCAGGCCCAAGCCCAACTCG 3718  
 Qy 1160 GluGlyIlePro-----IleMetArgTyrLeuSerArgGlnArgAsnSerLeu 1175  
 Db 3719 GAGGACCTGACCTCTGCAACCAACATCTGAAGTGAATCAGAAAGCAGAAATGCCAG 3778  
 Qy 1176 GlyGlyPheAlaSerThrGlnAspThrThrValAlaLeuLysValAlaLeuSerGluPheAla 1195  
 Db 3779 GGGGCTTCTCTCCACCGACAGACAGTGGTGGCTCTTCATCTCTGTCCAATATGGA 3838  
 Qy 1196 AlaLeuMetAsnThrGluArgThrAsnIleGlnValThrValThrGlyProSerSerPro 1215  
 Db 3839 CGCGCCACATTTAC---AGGACTGGAGAGGCTGCACAGGTACTATCCAGTCTTCAGGG 3895  
 Qy 1216 SerProLeuAlaValValGln----- 1222  
 Db 3896 ACAATTTCCAGAAATTCGAAGTGAACAACAATGCCCTGTACTGCAGAGTCTCA 3955  
 Qy 1223 -----ProMetAlaValAsnIleSerAlaAsnGlyPheGlyPheAlaIleCys 1238  
 Db 3956 TTGCGAGAGCTGCTGGGAAATACAGCATGAAGTGAACAGAGAAAGATGTCTTACTTC 4015  
 Qy 1239 GluLeuAsnValValTyrAsnValLysAlaSerGlySerSerArgArgArgSerIle 1258  
 Db 4016 CAGACCTCTTGAATTAACAATATT-----CTC 4042  
 Qy 1259 GlnAsnGlnIleAlaPheAspLeuAspValAlaValLys-----GluAsnLysAsp 1275

```

Db      4043 CCAAGAAAGAAAGTCCCTTGGCTTAGAGTCAGACTCTGCTCAAACTTGAT 4102
      1276 Aspleuasn---HisValaspleuasnValCysThrSerPheSerGlyProGly----- 1292
Db      4103 GAACCAAGAACCCACACGACGCTTCCAAATCTCCCTTAAGTGTCAATTACACAGAGACCCG 4162
      1293 ---ArgSerGlyMetAlaLeuMetGluValAsnLeuLeuSerGlyPheMetValProser 1311
Db      4163 TCTGCTCCCAACATGGCGCATGCTTAAGTGAAGATGCTCTGCTTCAATCCCGTAGAG 4222
      1312 GluAlaIleSerLeu-----SerGluThrValIleValIleValGluTyrAspHisGly 1328
Db      4223 CCAACAGTAAATGCTTGAAGATCTAACATGTCAGCGACGACAGCAAC 4282
      1329 LysLeuAsnLeuTyrIleLeuAspSerValAsnGluThrGlnPheCysValAsnIleProAla 1348
Db      4283 CATGCTTCAATTAACCTTGAATAGGAGTGCATAATCAGACACTGAGCTTGTTCACGGATT 4342
      1349 ValArgAsnPhelYsValSerAsnThrGlnAspAlaSerValSerIleValAspTyrTyr 1368
Db      4343 CTGCAAGATGTCCTCCAGTAGAGATCTCAACACGACATAGTAGAAGTCTATGATTAATC 4402
      1369 GluProArgArgGlnAlaValArgSerTyrAsnSer 1380
      4403 GAGACGATGAGTTGCAATCGCTGAGTACAAATGCT 4438

RESULT 5
US-09-873-403-4
/ Sequence 4, Application US/09873403
/ Patent No. US20020028207A1
/ GENERAL INFORMATION:
/ APPLICANT: SRIvesava, Pramod K
/ TITLE OF INVENTION: COMPLEXES OF ALPHA (2) MACROGLOBULIN AND ANTIGENIC
/ TITLE OF INVENTION: MOLECULES FOR IMMUNOTHERAPY
/ FILE REFERENCE: 8449-178
/ CURRENT APPLICATION NUMBER: US/09/873,403
/ CURRENT FILING DATE: 2001-06-04
/ PRIOR APPLICATION NUMBER: 09/625,139
/ PRIOR FILING DATE: 2000-07-25
/ PRIOR APPLICATION NUMBER: 60/209,266
/ PRIOR FILING DATE: 2000-06-02
/ SOFTWARE: FastSeq for Windows Version 3.0
/ SEQ ID NO 4
/ LENGTH: 4422
/ TYPE: DNA
/ ORGANISM: Homo sapiens
US-09-873-403-4

Alignment Scores:
Pred. No.: 3,36e-133 Length: 4422
Score: 1440.00 Matches: 451
Percent Similarity: 46.204 Conservative: 266
Best Local Similarity: 29.064 Mismatches: 551
Query Match: 19.604 Indels: 284
DB: 9 Gaps: 59

US-10-020-095-4 (1-1428) x US-09-873-403-4 (1-4422)
QY      12 LeuLeuCyValCysThrAlaIleValAlaValAlaProGlyProArgPheLeuValThr 31
      43 CTCCTTGCTCTCTCTGCCACAGCGCTCAGTCTCTGGAACCGAGATATATGTTCTG 102
Db      32 AlaProGlyIleIleArgProGlyGlyAsnValThrIleGlyValGluLeuLeuGlnHis 51
      103 GTCCCTCTCTCTCTCCACACT---GAGACCACTGAAGAGGCTGTGCTCTTCTGAGCTAC 159
QY      52 CysProSerGlnValThrValIleValAlaGluLeuLeuValThrAlaSerAsnLeuThrVal 71
Db      160 CTGAATGACAGACGTGCTGAAGTCTCTTGAAGTCTGTCAAGGGAACAGAGACCTTC 219
QY      72 ---SerValLeuGluAlaGluGly---ValPheGluIleGlySerPheIleThrLeuThr 89

```

```

Db      220 TTCACTGACCTGAGACCGAGATGACGTAATCCACTGTGTGCGCTTC----- 267
QY      90 LeuProSerLeuProLeuAsnSerAlaAspGlu-----IleTyrGluLeuArgVal 106
Db      268 ---GCTGTCCAAAGCTTCAATCAATGAGAGGATTAAGTTCTCTCACTGTCAAGAG 321
QY      107 ThrGlyArgThrGlnAspGluIleLeuPheSerAsnSerThrArgLeuSerPheGluThr 126
Db      322 AAAGACCAACCCCAAGAA-----TTTAAGAAGCGGACCAAGTGAATGTTAAGAAC 372
QY      127 LysArgIleSerValPheIleGlnThrAspIleValAlaLeuTyrLysProLysGlnGluVal 146
Db      373 GAGGACAGTGTGCTTGTTCAGACAGACAAATCAATCAATCAAAACAGGAGCAGACAGT 432
QY      147 LysPheArgIleValThrLeuPheSerAspPheLysProTyrLysThrSerLeuAsnIle 166
Db      433 AAATTTCTGTGTCTTCATGATGAATACTTCAACCCCTGAATGAGTTGATTCACCTA 492
QY      167 Leu---IleLysAspProLysSerAsnLeuIleGlnGlnThrPleuSerGlnIleSerAsp 185
Db      493 GTATACATTCACAGATCCCAAGAGAAATCGCATGCAATGCGACAGATTTCCAGTTAAG 552
QY      186 LeuGlyValIleSerLysThrPheGluLeuSerSerHisProIleLeuGlyAspTyrSer 205
Db      553 GGTGCGCTCAAGCAATTTCTTCTTCCCTCATGAGCCCTTCAGAGGCTCCTAACAAG 612
QY      206 IleGlnValGlnValAsnAsp-----GlnThrTyrTyrGlnSerPheGlnValSerGlu 223
Db      613 GTGTGTGTCAGACAGAAATCAGGTGAGAGCAGAGACAC---CCTTTACCCCTGAGAGAA 669
QY      224 TyrValLeuProLysPheGluValThrLeuGlnThrProLeuTyrCysSerMetAsnSer 243
Db      670 TTGTGTTCTTCCCAATTTGAAGTACAGTAACAGTCCCAAGATATATCAATCTTGGAA 729
QY      244 LysHisLeuAsnGlyThrIleThrAlaLysTyrThrTyrGlyLysProValLysGlyAsp 263
Db      720 GAAGAGATGAATGATCATGAGTGTGCGCTATACATATGGGAACGCTGCTGTGACAT 789
QY      264 ValThrLeuThrPheLeuProLeu----- 271
Db      790 GTGACTGTGAGCATTTGCGAAGATATATGACGCTTCCGACCTGCAGCGTGAAGATTCA 849
QY      272 -----SerPheThrGlyLys 276
Db      850 CAGGCTTTCTGTGAATAATTACGTGACAGCTAAACAGCCATGCGTCTTATATCAGCA 909
QY      277 LysLysAsnIleThrLysThrPheLysIleAsnGlySerAlaAsnPheSerPheAsnAsp 296
Db      910 GTMAAA-----ACCAAGTCTTCCAGCTGAAGAG-----AAGAG 945
QY      297 GluGluMetLysAsnValMetAspSerSerAsnGlyLysSerGlyTyrLeuAspLeuSer 316
Db      946 TATGAATGAAGAA-----CTTCACTAGAGGCCAGATCAAA 981
QY      317 SerProGlyProValGluIleLeuThr-----ThrValThrGluSerVal 331
Db      982 GAAGAGGAACAGGTGTGAANTGACTGAAGGACGTCCAGTGAATATCAAGAAACCAAT 1041
QY      332 ThrGlyIleSerArgAsnValSerThrAsnValPhePheLysGlnHisAspTyrIleIle 351
Db      1042 ACCAAACCTCA---TTTGTGAAGTGAATCAGCTTTCACAGAGG-----ATT 1089
QY      352 GluPhePheAspTyrThrThrValLeuLysProSerLeuAsnPheThrAlaThrValLys 371
Db      1090 CCTTCTTTGGG-----CAG 1104
QY      372 ValThrArgAlaAspGlyAsnGlnLeuThrLeuGluGluArgAsnAsnValValIle 391
Db      1105 GTGCGCTTGAAGATGGGAAAGCGTCCCTATACCA-----AATTAAGTCAATA 1152
QY      392 ThrValThrGlnArgAsnTyrThrGluTyrTrpSerGlySerAsnSerGlyAsnGlnLys 411

```

Db 1153 TTCATC---AGAGAAATGAAGCAAACTATTAC-----TCCAAATGCTACCAAGGATGAG 1203  
 Qy 412 MetGluAlaValGln---Lys1LeuSerThr-----ValProGlnSerGlyThrPhe 428  
 Db 1204 CATGGCCCTTGACGTTCTCTATCAACCAACCAAGCTTAAAGGTAAGTCTCTTAACTGTT 1263  
 Qy 429 Lys1LeuGluPhePro-----Lys1LeuGluAspSer 438  
 Db 1264 AGGGTCATATTACAAAGATCGTAGTCCCTGTTACAGGCTACAGAGGAGTGTCCAGAAACAC 1323  
 Qy 439 SerGluLeuGlnLeuLysAlaThrPheLeuGlySerLysSerSerMetAlaValHisSer 458  
 Db 1324 GAAGAGCAATCACTCACTGCTTAT----- 1347  
 Qy 459 LeuPheLysSerProSerLysThrTyrlLeuLeuLysThrArgAspGluAsn1LeuLys 478  
 Db 1348 CTTGTGTTCTCCCAAGCAAGAGCTTTGTCCACTTGAGCCCAATGCTCATGAATCAACC 1407  
 Qy 479 ValGlySer-----ProPheGluLeuValSerGlyAsnLys-----Arg 492  
 Db 1408 TGTGGCCATCTACAGACAGTCCAGGACATTAATTTGATGAGGAGGACCTGCTGGGG 1467  
 Qy 493 LeuLysGluLeuSer-----TyrMetValValSerArgGlyGlnLeuValAlaValGly 510  
 Db 1468 CTGAAGAGCTCTCTTTATTAATCTGATTAATGGCAAGGAGGAGGAGGAGGAGGAGGAG 1527  
 Qy 511 -----LysGlnAsnSerThr-----MetPheSerLeuThr 520  
 Db 1528 ACTGATGACCTGCTTGACAGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1587  
 Qy 521 ProGluAsnSerTrpThrProLysAlaCysVal1LeuValTyrlLeuLysAspGly 540  
 Db 1588 GTGAAGTACAGATGCTCTGTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1647  
 Qy 541 Glu1Leu1LeuSerAspValLeuLys1LeuProValGlnLeuValPheLysAsnLys1LeuLys 560  
 Db 1648 GACGTGATTTGGGGATTCTGCAAAATATGATGTTGAAATGCTGCGCCCAACAGAGTAT 1707  
 Qy 561 LeuTyrlTrpSerLysValLysAlaGluProSerGluLysValSerLeuArg1LeuSerVal 580  
 Db 1708 TTGAGCTTACAGCCATCAACAAAGTCTCCAGCTTCACACGCCCACTGCGAGTCAACAGC 1767  
 Qy 581 ThrGlnProAspSer1LeuValGly1LeuValAlaValAspLysSerValAsnLeuMetAsn 600  
 Db 1768 GCT---CCTAGCTCGTGTGGCCCTCGTCTGTGTGAGCAAAAGGAGTGTCTCATGAG 1824  
 Qy 601 AlaSerAsnAsp1LeuMetGluAsnValValHisGluLeu---GluLeuTyrlAsnThr 619  
 Db 1825 CCTGATGCTGAGCTCTCGGCGCTCGCTGTTTACAACTGCTACAGAAAGAGACCTCACT 1884  
 Qy 620 GlyTyrl-----TyrLeuGlyMet 625  
 Db 1885 GGCCTTCCCTGGGCTTTGAATGACCAAGAGAGAGATGACATGATCATCATGCTATATATCTC 1944  
 Qy 626 PheMetAsnSer-----PheAla 631  
 Db 1945 TATATTATGGAATCACATATATCTCACTATCAAGTACAAATGAAGAGATATGTACAGC 2004  
 Qy 632 ValPheGlnGlyCysGlyLeuTrpValLeuThrAspAlaAsnLeuThrLysAspTyrlLe 651  
 Db 2005 TTCTAGAGGAGCATGGCTTAAAGGCAATTCACCAACTCAAGATTCGTTAAACCCCAAAAT 2064  
 Qy 652 AspGlyValTyrlAspAsnAlaGlu-----Tyr 660  
 Db 2065 TGTCCACAGCTTCAACAGTATGAATGATGAGACCTGAGAGTCTAAGCTGAGTTTAT 2124  
 Qy 661 AlaGluTrpPheMetGluGluAsnGluGlyHis1LeuValAsp1LeuHisAspPheSerLeu 680  
 Db 2125 GAGTCAAGATGTAATGGAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2175  
 Qy 681 GlySerSerProHis-----ValArgLysHisPheProGluThrTrp1LeuTrpLeu 697  
 Db 2176 -----CCTCACACGGAAGACGTTAGAAAGTACTTCCCTGAGACATGTGATCTGGGAT 2226

Qy 698 AspThrAsnMetGlyTyrlArg1LeuTyrlGlnGluPheGluValThrValProAspSer1Le 717  
 Db 2227 TTGGTGTGTGTAACATCAGACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2286  
 Qy 718 ThrSerTrpValAlaThrGlyPheVal1LeuSerGluAspLeuGlyLeuGlyLeuThrThr 737  
 Db 2287 ACCGAGTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2346  
 Qy 738 ThrProValGluLeuGlnAlaPheGlnProPhe1LeuPheLeuAsnLeuProTyrlSer 757  
 Db 2347 ACT---GCTCTCTCCAGAGCTTCAGGCTCTTGTGTGAGGAGGAGGAGGAGGAGGAGGAGGAG 2403  
 Qy 758 Val1LeuArg1GlyGluGluPheAlaLeuGlu1LeuThr1LeuPheAsnTyrlLeuLysAspAla 777  
 Db 2404 GTGATTTGTGGAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2463  
 Qy 778 ThrGluValLysVal1LeuGlyLysSerAspLysPheAsp1LeuMetThrSerSer 797  
 Db 2464 ATCCGGGCTGAGTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2523  
 Qy 798 Glu-----LysAsnAlaThrGlyHisGlnGlnThrLeuLeuValProSer 812  
 Db 2524 GAACAGCGGCTCACTGATCTGTGCAACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2582  
 Qy 813 GluAspGlyAlaThrValLeuPhePro1LeuArgProThrHisLeuGlyGlu1LeuPro1Le 832  
 Db 2563 -----ACTGTGCTCTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2610  
 Qy 833 ThrValThrAlaLeuSer-----ProThrAla 841  
 Db 2611 ACTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2670  
 Qy 842 -----SerAspAlaValThrGluMet1LeuValLysValAlaGluGlyLe 856  
 Db 2671 CTGAACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2730  
 Qy 857 GluLysSerTyrlSerGlnSer1LeuLeuAspLeuThrAspAsnArgLeuGlnSerThr 876  
 Db 2731 GAGAGGAAACAAACATTCACATCCCTTCTTGTTCATCAGTGTGAGGTTCT----- 2784  
 Qy 877 LeuLysThrLeuSerPheSerPheProProAsnThrValThrGlySerGluArgValGln 896  
 Db 2785 ---GAAGAAATTTCCCGAAGATGCTCCCAAAAGTGTGTAAGAAATGCTGCCGCGCTTCT 2841  
 Qy 897 1LeuThrAla1LeuGlyAspValLeuGlyProSer1LeuAsnGlyLeuAlaSerLeu1LeuArg 916  
 Db 2842 GTCTCAGTTTGGAGACATTAATGAGCTGTGCAATGCAAAACACACAAAAATCTTCTCAG 2901  
 Qy 917 MetProTyrlGlyCysGlyGluGlnAsnMet1LeuAsnPheAlaProAsn1LeuTyrlLeu 936  
 Db 2902 ATGCCCTATGAGCTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2961  
 Qy 937 AspTyrlLeuThrLysLysGlyLeuLeuThrAspAsnLeuLysGlyValAlaLeuSerPhe 956  
 Db 2962 GATTATCTTAATGAACACAGAGCTTATCTCAGAGAGTCAAGTCAAGGAGGAGGAGGAGGAGGAG 3021  
 Qy 957 MetArgGlnGlyTyrlGlnArgGluLeuLeuTyrlGlnArgGluAspGlySerPheSerAla 976  
 Db 3022 CTCAACACTGGTTACAGAGACAGTGAATCAACAAACATATATGAGTCTCTCAACAGACCC 3081  
 Qy 977 PheGly---AsnTyrlAspProSer---GlySerThrTrpLeuSerAlaPheValLeuArg 994  
 Db 3082 TTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3141  
 Qy 995 CysPheLeuGluAlaAspProTyrlLeuAsp1LeuAspGlnAsnValLeuHisArgThrTyrl 1014  
 Db 3142 ACTTTGCCAGAGCTGAGGCTCATCTCATGATGAACACACATTAACCAAGCCCTC 3201  
 Qy 1015 ThrTrpLeuLysGlyHisGlnLysSerAsnGlyGluPheTrpAspProGlyArgVal1Le 1034  
 Db 3202 ATATGGCTTCCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3261



Oy	1035	HisSerGluLeuGlnGlyGlyAsnLysSerProValThrLeuThrAlaTyrIleValThr	1054
Db	3262	AACATGCCATTAAGGAGGAGTAGTAAGAAATGAAGACCCCTCCCGCTTAATACCAATC	3321
Oy	1055	SerLeuLeuGlyTyrArgLysTyrGlnProAsnIleAspValGlnGluSerIleHisPhe	1074
Db	3322	GCCCTTCGGAGATTCCTCTCAACAGTCACTACCCCTGTGTGTCCGAATGCCCTGTTTGC	3381
Oy	1075	LeuGluSer-----GluPheSerArgGlyIleSerAspAsnTyrThr	1088
Db	3382	CTGGAGTCAGCTGGAGACAGCAACAAGGGGACCATGGC---AGCAGATGATATATCC	3438
Oy	1089	LeuAlaLeuIleThrTyrAlaLeuSerSerValGlySer---ProValAlaGlyGluAla	1107
Db	3439	AAAGCATGCTGGCCTATGCTTGTTCCTCGCAGGTATACAGGACAAGAGAGAAAGTA	3498
Oy	1108	LeuAsnMetLeuThrTyrArgAlaGluGlnGlyGlyMetGlnPheTrp-----	1124
Db	3499	CTCAAGTCACTTAATGAGAAAGCTGTGAAGAAAGACAACTGTTCATTGGAGAGCCCT	3558
Oy	1125	ValSerSerGluSerLysLysSerAspSerTrpGlnProArg-----SerLeuAspIle	1142
Db	3559	CAGAAACCCAAAGGACCAAGTGGGGGATTTTACAAACCCAGGCTCCCTCTCGAGAGTG	3618
Oy	1143	GluValAlaAlaTyrAlaLeuLeuSerHisPheLeuGlnPheGln-----ThrSer	1159
Db	3619	GAGATGCATCTATGTCCTCTCGCT---TATCTCAGCGCCAGCCAGCCCAACTCG	3675
Oy	1160	GluGlyIlePro-----IleMetArgTrpLeuSerArgGlnArgAsnSerLeu	1175
Db	3676	GAGGACCTGACCTCTGCAACCAACTCTGTGAAGTGATCAGAAAGACAGCAAAATGCCAG	3735
Oy	1176	GlyGlyPheAlaSerThrGlnAspThrThrValAlaLeuLysAlaLeuSerGlnPheAla	1195
Db	3736	GGCGGTTCTCTCCACCAACGACAGCAGTGTGTGCTCCATGCTCTGTCCAAATATGCA	3795
Oy	1196	AlaLeuMetAsnThrGlnArgThrAsnIleGlnValThrValThrGlyProSerSerPro	1215
Db	3796	GCCGCGCATTTTACC---AGGACGTGGAGGCTGCACAGTGAATCACTCACTTCAGAGG	3852
Oy	1216	SerProMetAlaValValGln-----	1222
Db	3853	ACATTTTCCACCAAAATTCCAAGTGACAAACAATCGCTGTACTGCAGCAGGTCTCA	3912
Oy	1223	-----PrometAlaValAsnIleSerAlaGlnGlyPheGlyPheAlaIleCys	1238
Db	3913	TTGCCAAGACGTGCTCGTGGGAATACAGCATGAAAGTACAGAGAAAGATGTGTCTACCTC	3972
Oy	1239	GlnLeuAsnValValTyrAsnValLysAlaSerGlySerSerArgArgArgSerIle	1258
Db	3973	CAGACCTCTCTGAATACATATAT-----CTC	3999
Oy	1259	GlnAsnGlnGluAlaPheAspLeuAspValAlaValLys-----GluAsnLysAsp	1275
Db	4000	CCAGAAAGGAAGGATTCCTCTTGTCTTTAGAGTGCAGACTGTGCTCTCAAACTGTGAT	4059
Oy	1276	AspLeuAsn---HisValAspLeuAsnValCysThrSerPheSerGlyProGly-----	1292
Db	4060	GAACCCAAAGCCACACCAAGCTTCCAAATCTCCCTTAAGTCAAGTACACAGGAGCCGC	4119
Oy	1293	---ArgSerGlyMetAlaLeuMetGlnValAsnLeuLeuSerGlyPheMetValProSer	1311
Db	4120	TCTGCTCTCAACATGGCGATCGTTGATGTGAAGATGGCTCTGGCTTCATTCCTCGTAG	4179
Oy	1312	GluAlaIleSerLeu-----SerGluThrValLysValGluTyrAspHisGly	1328
Db	4180	CCAAACAGGAAAGATGCTTGAAGAATCTAACCATGTGACCCGAGACAGTACAGCAAC	4239
Oy	1329	LysLeuAsnLeuTyrLeuAspSerValAsnGluThrGlnPheCysValAsnIleProAla	1348
Db	4240	CATGCTCTGATTTTACCTTGTATAGGTGTCAAAATCAGACACTGAGCTTGTCTTCAAGTT	4299
Oy	1349	ValArgAsnPheLysValSerAsnThrGlnAspAlaSerValSerIleValAspTyrTyr	1368

Dd		4300	CTGCAAGATGTCCAGTAAAGACATCTCAAAACCGCATAGTAAGTAATTACTAC	4355
Oy		1369	GluProAArgGIlnAlaValArgSerTYrAsnSer	1380
Dd		4360	GAGACGGATGAAGTTTGCAATGCCCTGAGTACAATGCT	4395
 RESULT 6 US-09-873-403-3 ; Sequence 3, Application US/09873403 ; Patent No. US20020028207A1 GENERAL INFORMATION: APPLICANT: Srivastava, Pramod K TITLE OF INVENTION: COMPLEXES OF ALPHA (2) MACROGLOBULIN AND ANTIGENIC FILE REFERENCE: 8449-178 CURRENT FILING DATE: 2001-06-04 PRIOR APPLICATION NUMBER: 09/625,139 PRIOR FILING DATE: 2000-07-25 PRIOR APPLICATION NUMBER: 60/209,266 PRIOR FILING DATE: 2000-06-02 NUMBER OF SEQ ID NOS: 5 SOFTWARE: FastSeq for Windows Version 3.0 SEQ ID NO 3 LENGTH: 4577 TYPE: DNA ORGANISM: Homo sapiens US-09-873-403-3				
 Alignment Scores:				
Pred. No.:	3,56e-133	Length:	4577	
Score:	1440.00	Matches:	451	
Percent Similarity:	46.20%	Conservative:	266	
Best Local Similarity:	29.06%	Mismatches:	551	
Query Match:	19.60%	Indels:	284	
DB:	9	Gaps:	59	
 US-10-020-095-4 (1-1428) x US-09-873-403-3 (1-4577)				
Oy		12	LeuleuCysValCysThrAlaIalaLeuAlaValAlaProGlyProArgPheLeuValThr	31
Dd		86	CTCTGTGCCTCTCTCCCAACAGCCGCTCAGCTCTCGAAAAACCGAGATAATAGTTCG	145
Oy		32	AlaProGlyIleIleIleArgProGlyGlyAsnValThrIleGlyValGluLeuGluHis	51
Dd		146	GTCCCCCTCCCTGCTCACACT--GAGACCCTAGAGAGGGCGTGTCTCTTGACCTAC	202
Oy		52	CysProSerGlnValThrValLysAlaGluLeuLeuLysThrAlaSerAsnLeuThrVal	71
Dd		203	CTGAATGAGAACAAGTACTGTAGTAGTCTCTTGAGTCTGTCAAGGGAAAACGAGACCTC	262
Oy		72	--SerValLeuGluAlaGluGly--ValPheGluLysGlySerPheLysThrLeuThr	89
Dd		263	TTCACCTGACCTGGAGCGGAGAGATGACCTACTCCACTGTGTGCCTTC-----	310
Oy		90	LeuProSerLeuProLeuAsnSerAlaAspGlu-----IleTyrlGluLeuArgVal	106
Dd		311	----GCTGTCCCAAAGCTTCAATCCAAATGAGAGATATGTTCTCTCACTGCCAAGTG	364
Oy		107	ThrgIyArGthrGlnAspGluIleLeuPheSerAsnSerThrArgLeuSerPheGluThr	126
Dd		365	AAAGAACCAACCAAGAA-----TTTAAGAAAGCGGACCAACAGTATGCTTAAGAAC	415
Oy		127	LysbArgIleSerValPheIleGlnThrAspLysAlaLeuTyrlLysProLysGlnGluVal	146
Dd		416	GAGGACAGCTGTGCTCTTGTCCAGACAGCAATCAATCTCAAAACGAGGCGACAGATG	475
Oy		147	LysPheArgIleValThrLeuPheSerAsnPheLysProTyrlLysThrSerLeuAsnIle	166
Dd		476	AAATTTCGTTGTCTCCATGGATGAATAAACTTCAACCCCTGAATGATGATTCACTA	535
Oy		167	Leu---IleLysAspProLysSerAsnLeuIleGlnGlnTrpLeuSerGlnGlnSerAsp	185

```

Db      536 GTATACATTCAGATCCCAAGAAATGCATCGCACATGCGACAGTTTCAGTTACAG 595
      186 LeuGlyValIleSerLeuThrPheGlnLeuSerSerHisProIleLeuGlyAspTrpSer 205
      596 GGTGGCTCCAGAACATTTTCTTTCTTCCCTCATAGAGCCCTTCCAGGGCTCCACAG 655
      206 IleGlnValGlnValAsnAsp-----GlnThrTyTyTyGlnSerPheGlnValSerGlu 223
      656 GTGGGTGACAGAGAAATCAGGTGGAGAGACAGACAC-----CCTTACCGTGGAGAA 712
      224 TyrValIleuProLysPheGlnValThrLeuGlnThrProLeuTyTySerSerMetAsnSer 243
      713 TTTGTTCTTCCCAAGTTGAAAGTACAGTAAACAGTCCCAAGATGATACATCTTGGAA 772
      244 LysHisLeuAsnGlyThrIleThrAlaLeuTyTyTyTyTyTyTyTyTyTyTyTyTyTyTyTy 263
      773 GAAGAGATGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 832
      264 ValThrLeuThrPheLeuProLeu----- 271
      833 GTGACTGTGACATTTGGAGAAAGTATAGTACGCTTCCGACCTGCCAGCGTAAGATTCA 892
      272 -----SerPheTrpGlyLys 276
      893 CAGGCTTTCTGTGAGAAATTCAGTGCAGACGTAACAGCCATGCGCTGCTTATCAGCAA 952
      277 LysLysAsnIleThrLeuThrPheLysIleAsnGlySerAlaAsnPheSerPheAsnAsp 296
      953 GTAAAA-----ACCAGAGCTCTCCAGCTGAAAGAG-----AAGAG 988
      297 GluGluMetLysAsnValMetAspSerSerAsnGlyLeuSerGlyTyTyLeuAspLeuSer 316
      989 TATGAAATGAAA-----CTTCACACTGAGGCCACAGATCCAA 1024
      317 SerProGlyProValGlnIleLeuThr-----ThrValThrGlnSerVal 331
      1025 GAAGAGAGAGACAGTGGTGAATTGACTGAAAGGACAGTCCAGTGAATACAAAGAACATA 1084
      332 ThrGlyIleSerArgAsnValSerThrAsnValPhePheLysGlnHisAspTyTyle 351
      1085 ACCAAACTCTCA-----TTTGTAAAGTGCATCTCACTTGCACAGGAA-----ATT 1132
      352 GluPhePheAspTyTyThrValLeuLysProSerLeuAsnPheThrAlaThrValLys 371
      1133 CCCTTCTTGGG-----CAG 1147
      372 ValThrArgAlaAspGlyAsnGlnLeuThrLeuGlnLysArgAsnValValIle 391
      1148 GTGGCGCTAGTAGATGGAAAGCGCTCCCTATACCA-----AATAAGTCATA 1195
      392 ThrValThrGlnArgAsnTyTyThrGlyTyTyTrpSerGlySerAsnSerGlyAsnGlnLys 411
      1196 TTCATC-----AAGAGAAATGAAGCAAACTATTAC-----TCCAAATGCTACCAACGAGTATG 426
      412 MetGlnAlaValGln-----LysIleAsnTyTyThr-----ValProGlnSerGlyTyThrPhe 428
      1247 CATGGCTTGTACAGTCTCTATACACACACACCAACGTTATGGTACTCTTAACTTT 1306
      429 LysIleGluPhePro-----IleLeuGlnLysSer 438
      1307 AGGGTCATTTACAGAGATCGTATCCCTGTTACGGCTACACAGTGGGTGTCAAGAAACAC 1366
      439 SerGluLeuGlnLeuLysAlaTyTyPheLeuGlySerLysSerSerMetAlaValHisSer 458
      1367 GAAGAGGACATCACACTCTTAT----- 1390
      459 LeuPheLysSerProSerLysThrTyTyleGlnLeuLysThrArgAspLysAsnIleLys 478
      1391 CTGTGTTCTCCCAAGCAAGAGCTTGTCCACTTGGCCCAATGTCATGAACCTACCC 1450
      479 ValGlySer-----ProPheGlnLeuValValSerGlyAsnLys-----Arg 492

```

```

Db      1451 TGTGGCATACTCAGACAGTCCAGGACCATTAATTCTGAATGAGGACCCCTGCTGGG 1510
      493 LeuLysGluLeuSer-----TyrMetValValSerArgGlyGlnLeuValAlaValGly 510
      1511 CTGAAGAAGCTCTCTTTTATATCTGATATGCAAAAGGAGGACATTTGCCAAGCTGGG 1570
      511 -----LysGlnAsnSerThr-----MetPheSerLeuThr 520
      1571 ACTCATGAGACTGTTGGAAGCAGAGAGACATGAAGGCCATTTCATCTCAATCCCT 1630
      521 ProGlnAsnSerTrpThrProLysAlaCysValIleValTyTyTyTyleGlnAspArgGly 540
      1631 GTGAAGTCAACATTCCTCTGTCTGCTGGTGTCTCATCTATGCTGTTTAACTTACCGGG 1650
      541 GluIleIleSerAspValIleLysIleProValGlnLeuValPheLysAsnLysIleLys 560
      1691 GACCTGATTTGGGATTTCTGCAAAATATGATGATTTGAAAATTGTCGGCCACCAAGTGGAT 1750
      561 LeuTyTyTrpSerLysValLysAlaGluProSerGlyLysValSerLeuArgIleSerVal 580
      1751 TTGAGCTTACAGCCCATCACAAGTCTCCAGCTCCACAGCCCACTGGCAGTACAGCG 1810
      581 ThrGlnProAspSerIleValGlyIleValAlaValAspLysSerValAsnLeuMetAsn 600
      1811 GCT-----CCTCAGTCCGCTGCGCCCTCGCTGTGTGGACCAAGCGTGTCTCATTAAG 1867
      601 AlaSerAsnAspIleThrMetGlnAsnValValHisGluLeu-----GluLeuTyTyAsnThr 619
      1868 CTGTATCTGAGCTCTGGGCTCTCGGTTTACCACTGCTTACCAAAAAGAGACTCATCT 1927
      620 GlyTyTy-----TyrLeuGlyMet 625
      1928 GGTTCCTCGGCGCTTGAATGACACAGACGATGAAGACTGCATCAATCTCATTAATGTC 1987
      626 PheMetAsnSer-----PheAla 631
      1988 TATATTATATGAATCATATATCTCCAGTATCAAGTACAATGAAGATATGATGACG 2047
      632 ValPheGlnGlyCysGlyLeuTyTyValLeuThrAspAlaAsnLeuThrLysAspTyTyle 651
      2048 TTCTTAGAGACATGGCTTTAAAGCATTCACCACTCAAAAGATTCTTAAACCAAAATG 2107
      652 AspGlyValTyTyAspAlaVal-----Tyr 660
      2108 TGTCCACAGCTTCAACAGTATGAATGATCATGACCTGAGGTCTTACGTAGTATTAT 2167
      661 AlaGluArgPheMetGluAsnGlyHisIleValAspIleHisAspPheSerLeu 680
      2168 GAGTCAGATGTAATGGAGAGGCCATGACAGCTGTGTGATGTTGAAG----- 2218
      681 GlySerSerProHis-----ValArgLysHisAspPheProGluTyTyTrpIleTrpLeu 697
      2219 -----CCTCACAGGAGACCGTACGAAAGTACTTCCCTGAGACATGATTTGGAT 2269
      698 AspThrAsnMetGlyTyTyArgIleTyTyGlnLysPheGlnValValThrValProAspSerIle 717
      2270 TTGGTGTGTGAATACTAGACAGAGGTGGCTGAGGTAGAGATTAACAGTCCCGACACATC 2329
      718 ThrSerTyValAlaThrGlyPheValIleSerGluAspLeuGlyLeuGlyLeuThr 737
      2330 ACCCAGTGAAGAGGAGGCGCTTCTGCTGTGAAATCTGAGACTGTGATCTTCTCC 2369
      738 ThrProValGluLeuGlnAlaPheGlnProPhePheIlePheLeuAsnLeuProTyTySer 757
      2390 ACT-----GCTCTCTCCAGGCTTTCAGCCCTTCTTGTGAGACTTACAAAGCCCTTACTCT 2446
      758 ValIleArgGlyGlnGluPheAlaLeuGlnIleThrIlePheAsnTyTyLeuLysAspAla 777
      2447 GTGATTTGTGAGAGAGGCTTACACTCAGACGACAGCTCTTAACCTTCCCAAAATGC 2506
      778 ThrGluValLysValIleIleGluLysSerAspLysPheAspIleLeuMetThrSerSer 797
      2507 ATCCGGTCAAGTGTGACGTGAAAGCTTCCGCGCTTCTTGTCTGCCAGTGAAG 2566

```



```

; PRIOR APPLICATION NUMBER: US 60/211,379
; PRIOR FILING DATE: 2000-06-14
; PRIOR APPLICATION NUMBER: US 60/237,054
; PRIOR FILING DATE: 2000-10-02
; NUMBER OF SEQ ID NOS: 3950
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 2236
; LENGTH: 4577
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. US20020142981A1 M11313
; US-09-880-107-2236

Alignment Scores:
Pred. No.: 3,566-133 Length: 4577
Score: 1440.00 Matches: 451
Percent Similarity: 46.20% Conservative: 266
Best Local Similarity: 29.06% Mismatches: 551
Query Match: 19.60% Indels: 284
DB: 10 Gaps: 59

US-10-020-095-4 (1-1428) x US-09-880-107-2236 (1-4577)
Qy 12 LeuLeuCyValCyThrAlaAlaValAlaProGlyProArgPheLeuValThr 31
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 86 CTCCTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 145
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 32 AlaProGlyIleIleArgProGlyGlyAsnValThrIleGlyValGluLeuGluHis 51
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 146 GTCCCTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 202
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 52 CysProSerGlnValThrValIleValIleLeuLeuLeuValThrAlaSerAsnLeuThrVal 71
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 203 CTGAATGACAGACGTGACTGTGAAGCTCTTCTGAGCTGTCAGGGGAAACAGAGCCTC 262
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 72 --SerValIleGluIleGluGly--ValPheGluYsgIleSerPheLeuThr 89
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 263 TTCACTGACCTGGAGGGCGGAGAAATGACGTACTCCAGCTGTCGCTTC----- 310
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 90 LeuProSerLeuProLeuAsnSerAlaAspGlu-----IleTyrGluLeuArgVal 106
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 311 -----GCTGTCCTCCAAAGCTTTCATTCATAGAGAGGTATGTTCTTCACTGTCAGGTG 364
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 107 ThrGlyArgThrGlnAspGluIleLeuPheSerAsnSerThrArgLeuSerPheGluThr 126
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 365 AAAGACCAACCCCAAGAA-----TTTAAAGACCGGACCAACAGTATGTTAAAGAC 415
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 127 LysArgIleSerValPheIleGlnThrAspLysAlaLeuTyrLysProLysGlnIleVal 146
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 416 GAGGACACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 475
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 147 LysPheArgIleValThrLeuPheSerAspPheLysProTyrLysThrSerLeuAsnIle 166
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 476 AAATTTCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 535
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 167 Leu---IleLysAspProLysSerAsnLeuIleGlnIleThrLeuSerGlnIleSerAsp 185
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 536 GTATATCATTCAGAGATCCAAAGAAATGCAATCGCACATGCGACAGATTTCCAGTTAGAG 595
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 186 LeuGlyValIleSerLysThrPheGlnLeuSerSerHisProIleLeuGlyAspTrpSer 205
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 596 GGTGGCCCTCAGCAATTTCTTTCCCTCTCATAGAGCCCTTCCAGGCTCTCTACAG 655
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 206 IleGlnValGlnValAsnAsp-----GlnThrTyrTyrGlnSerPheGlnValSerGlu 223
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 656 GTGGGTGACAGAAATCAGTGAAGGACAGAGCAC---CCTTTCAACCGTGAAGGAA 712
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 224 TyrValLeuProLysPheGlnValThrLeuGlnThrProLeuTyrArgSerMetAsnSer 243
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 713 TTTGTCTTCTCCCAAGTTTGAAGTACAGATGACAGTCCAAAGATATCATCATCTTGA 772
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 244 LysHisLeuAsnGlyThrIleThrAlaLysTyrThrTyrGlyLysProValIleGlyAsp 263

```

```

Db 773 GAAGAGATGATATATACAGTGTGCTGCTTATACATATGGGAAGCTGCTCCCTGACAT 832
Qy 264 ValThrLeuThrPheLeuProLeu----- 271
Db 833 GTGACTGTGACATTTGACAGAAATATAGACCTTCCGACTGCCACGGTGAAGATTCA 892
Qy 272 -----SerPheTrpGlyLys 276
Db 893 CAGGCTTCTGTGAGAAATTCAGTGGACAGCTAAACAGCATGCTGCTTCTATACAGCA 952
Qy 277 LysLysAsnIleThrLysThrPheLysIleAsnGlySerAlaAsnPheSerPheAsnAsp 296
Db 953 GTAATA-----ACCAAGGCTTCCAGCTGAAGAG-----AAGAG 988
Qy 297 GluGluMetLysAsnValMetLeuPheSerSerAsnGlyLeuSerGlyLysLeuAspLeuSer 316
Db 989 TATGAATGAAA-----CTTCACTGAGGCCCAAGATCCAA 1024
Qy 317 SerProGlyProValGluIleLeuThr-----ThrValThrGlnSerVal 331
Db 1025 GAAGAGGAAACAGTGTGGAATTTGACTGGAAGGACGTCCAGTGAATCACAAGAACATA 1084
Qy 332 ThrGlyIleSerArgAsnValSerThrAsnValPhePheLysGlnHisAspTyrIleIle 351
Db 1085 ACCAACTCTCA---TTTGTGAAGTGAAGTCACTCACTTTCAGACAGGGA-----ATT 1132
Qy 352 GluPhePheAspTyrThrThrValIleLysProSerLeuAsnSerThrAlaThrValLys 371
Db 1133 CCCTCTTTGGG-----CAG 1147
Qy 372 ValThrArgAlaAspGlyAsnGlnLeuThrLeuGluGluValArgAsnAsnValIle 391
Db 1148 GTGGCGCTTATATAGTGAAGAGCGTCCATACCA-----ATAAAGTCATA 1195
Qy 392 ThrValThrGlnArgAsnTyrThrGlyTyrTrpSerGlySerAsnSerGlyAsnGlnLys 411
Db 1196 TTCATC---AGAGAAATGAAAGCAAACTATTAC-----TCCATATCTACACAGATGAG 1246
Qy 412 MetGluAlaValGln---LysIleAsnTyrThr-----ValProGlnSerGlyThrPhe 428
Db 1247 CATGGCTTATACAGTTCTCTATCAACACCAAGATTATGGGTACTCTTCACTGTT 1306
Qy 429 LysIleGluPhePro-----IleLeuGluAspSer 438
Db 1307 AGGTCATATACAGAGATGCTAGTCCCTGTACGCGTACCAAGGGGTGTCAGAAAGAC 1366
Qy 439 SerGluLeuGlnLeuLysAlaTyrPheLeuGlySerLysSerSerMetAlaValHisSer 458
Db 1367 GAAGAGGACATCATCACTGCTAT----- 1390
Qy 459 LeuPheLysSerProSerLysThrTyrIleGlnLeuLysThrArgAspGluAsnIleLys 478
Db 1391 CTGTGTTCTCCCAAGCAAGACTTTGTTCACCTTGAAGCCATGTCTCATGAATCAACC 1450
Qy 479 ValGlySer-----ProPheGluLeuValIleSerGlyAsnLys-----Arg 492
Db 1451 TGTGGCATATCTAGACAGTCCAGGACATTAATTTCTGAATGAGGACACCTGCTGGGG 1510
Qy 493 LeuLysGluLeuSer-----TyrMetValIleSerArgGlyGlnLeuValAlaValGly 510
Db 1511 CTGAAGAGCTGCTCTTTATTTATCTGATATGCAAGAGGAGGACATTGCCAATCGGG 1570
Qy 511 -----LysGlnAsnSerThr-----MetPheSerLeuThr 520
Db 1571 ACTCATGACTGCTGTTGGAAGGAGGAAGACATGAAGGCGCATTTTCCATCAATTCCT 1630
Qy 521 ProGluAsnSerTrpThrProLysAlaCysValIleValTyrTyrIleGluAspGly 540
Db 1631 GTGAAGTCAACATTCCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1690
Qy 541 GluIleIleSerAspValLeuLysIleProValGlnLeuValPheLysAsnLysIleLys 560

```

Db 1691 GACGTCATTGGGATTCGCCAAATATGATGTTGAAATTTGCTGCGCAACAAGTGAT 1750  
 QY 561 LeuTYrTTPSerXyValYsAlaGluProSerGluLysValSerLeuArgIleSerVal 580  
 Db 1751 TTGAGCTTACGCCATCACAAGTCTCCAGCTTCACAGCCCACTGCGAGTCAAGG 1810  
 QY 581 ThrGlnProAspSerIleValGlyIleValAlaValAspLysSerValAsnLeuAsn 600  
 Db 1811 GCT---CCTCAGTCCCTGCGCCCTCCGCTGCTGTGACCAAAAGCGTGTCTCATGAAG 1867  
 QY 601 AlAserAsnAspIleThrMetGluAsnValValHisGluLeu---GluLeuTYrAsnThr 619  
 Db 1868 CCTGATGCTGAGCTCTGGCGCTCTCGTTTACAACTGCTACCAAGAAAGAAAGAACTCACT 1927  
 QY 620 GlyTYr-----TyrLeuGlyMet 625  
 Db 1928 GCGTTCCCGCGGCTTTGAATGACAGACGATGAAGACTGCATCATGTCTAATATGTC 1987  
 QY 626 PheMetAsnSer-----PheAla 631  
 Db 1988 TATATTAAATGATCATCATATACTCCAGTATCAATGACAAATGAAAGATATGTACAC 2047  
 QY 632 ValPheGlnGluCysGlyLeuTYrValLeuThrAspAlaAsnLeuThrIleAspTYrIle 651  
 Db 2048 TTCTTAAGAGACATGCGCTTAAAGGCAATTCACCACTCAAGATTGTTAAACCAAAATG 2107  
 QY 652 AspGlyValTYrAspAsnAlaGlu-----Tyr 660  
 Db 2108 TGTCCACAGCTTCAACAGTATGAAATGATGACCTGAAGGCTCTACGTTAGCTTTTAT 2167  
 QY 661 AlaGluArgPheMetGluGlnAsnGluGlyHisIleValAspIleHisAspPheSerLeu 680  
 Db 2168 GAGTCAGATGATATGAGAGAGAGCCATGACGCGCTGTGATGTAAGAG----- 2218  
 QY 681 GlySerSerProHis-----ValArgGlyHisPheProGluThrTrpIleTrpLeu 697  
 Db 2219 -----CCTCACACGAGACCGGTACGAAAGTATCTTCCGAGACATGATCTGGGAT 2269  
 QY 698 AspThrAsnMetGlyTYrArgIleTYrGlnGluPheGluValThrValProAspSerIle 717  
 Db 2270 TTGCTGTGTGTAACTCAGACAGGGCTGCTGAGTATGAGATACAGTCTCCAGACCAATC 2329  
 QY 718 ThrSerTrpValAlaThrGlyPheValIleSerGluAspLeuGlyLeuGlyLeuThr 737  
 Db 2330 ACCGATGAGAGGAGGCGGCTTCTGCTGCTGAGAGATCTGGAATCTGTATCTCTCC 2389  
 QY 738 ThrProValGluLeuGlnAlaPheGlnProPhePheIlePheLeuAsnLeuProTYrSer 757  
 Db 2390 ACT---GCCTCTCTCGAGCCTTCCAGCCCTTCTTGTGAGACTTTCATATGCTTACTCT 2446  
 QY 756 ValIleArgGlyGluGluPheAlaLeuGluIleThrIlePheAsnTYrLeuLysAspAla 777  
 Db 2447 GGTATTCGTGAGAGGCTTTCACACTCAAGGCGCAAGCTTAACCTTCCCAATGCG 2506  
 QY 778 ThrGluValLysValIleIleGluLysSerAspLysPheAspIleLeuMetThrSerSer 797  
 Db 2507 ATCCGGGTAGTGTGAGTGAAGAGCTCTCCCGCTTCTTGTCTGTCCAGTGAAGAG 2566  
 QY 798 Glu-----IleAsnAlaThrGlyHisGlnGlnThrLeuLeuValProSer 812  
 Db 2567 GAACAAGCCCTCATCTGCTGTGCAACGGGCGGCA----- 2605  
 QY 813 GluAspGlyAlaThrValLeuPheProIleArgProThrHisLeuGlyGluIleProIle 832  
 Db 2606 -----ACTGTCTCTGGGCGATTAACCCCAAGTCAATAGAAATGTGAATTC 2653  
 QY 833 ThrValThrAlaLeuSer-----ProThrAla 841  
 Db 2654 ACTGTAGGCGCAGAGCACTAAGTCTCAAGCTCTGTGGAGCTAGAGTGCCTTCAAGT 2713  
 QY 842 -----SerAspAlaValThrGlnMetIleLeuValLysAlaGluGlyIle 856  
 Db 2714 CCTGAACAGGAAGAAAGACACAGTCAATCAAGCTCTGTGTGTGAACCTGAAGAGACTA 2773

QY 857 GluLysSerTYrSerGlnSerIleLeuLeuAspLeuThrAspAsnArgLeuGlnSerThr 876  
 Db 2774 GAGAAAGAAACAACTTGAATCACTCCCTACTTTCATCAGAGTGGAGTTC----- 2827  
 QY 877 LeuLysThrLeuSerPheSerPheProAsnThrValThrGlySerGluArgValGln 896  
 Db 2828 ---GAAAGATTAATCCCTGAAACTGCCCAAAATGTGTGAAGATCTGCCGAGCTTCT 2884  
 QY 897 IleThrAlaIleGlyAspValLeuGlyProSerIleAsnGlyLeuAlaSerLeuIleArg 916  
 Db 2885 GTCTAGTTTGGAGACATATATAGCTCTGCCAAGCAAAACACAAATCTTCTCCAG 2944  
 QY 917 MetProTYrGlyCysGlyGluAsnMetIleAsnPheAlaProAsnIleTYrIleLeu 936  
 Db 2945 ATGCCCTATGCTGTGGAGAGACAAATATGCTCTTGTCTCTTAACATTAATGACTG 3004  
 QY 937 AspTYrLeuThrLysLysGlnLeuThrAspAsnLeuLysGluLysAlaLeuSerPhe 956  
 Db 3005 GATTATCTAAATGAACACAGAGCTTACTCCAGAGCTCAAGTCCAGGCCATTTGGCTAT 3064  
 QY 957 MetArgGlnGlyTYrGlnArgGluLeuLeuTYrGlnArgGluAspGlySerPheSerAla 976  
 Db 3065 CTCACACTGTGTACAGAGACAGTTGAACAAACATATGATGCTCTCCACAGCAC 3124  
 QY 977 PheGly---AsnTYrAspProSer---GlySerThrTrpLeuSerAlaPheValLeuArg 994  
 Db 3125 TTGGGAGGCGATATGAGAGAACAGGCGCAACCTGCTACAGCCTTGTGTGTGAAG 3184  
 QY 995 CysPheLeuGlnAlaAspProTYrIleAspIleAspGlnAsnValLeuHisArgThrTYr 1014  
 Db 3185 ACTTTTGCCCAAGCTCGAGCCTACATCTTCATGATGAAGACACATACCAAGCCCTC 3244  
 QY 1015 ThrTrpLeuLysGlyHisGlnLysSerAsnGlyGluPheTrpAspProGlyArgValIle 1034  
 Db 3245 ATATGCTCTCCACAGGACAAAGACATGCTGTTCAGAGAGCTGTGGCTCACTGCTC 3304  
 QY 1035 HisSerGluLeuGlnGlyLysAsnLysSerProValThrLeuThrAlaTYrIleValThr 1054  
 Db 3305 AACATGCCATTAAGAGAGAGAGTGAAGATGAAGTACCTCTCCGCTTAATTAACATC 3364  
 QY 1055 SerLeuLeuGlyTYrArgLysTYrGlnProAsnIleAspValGlnGluSerIleHisPhe 1074  
 Db 3365 GCCCTTCTGAGATTCCTCTCAGACTCACTCAACCTGTTGTCGGAATGCCCTGTTTGC 3424  
 QY 1075 LeuGlnSer-----GluPheSerArgIleIleSerAspAsnTYrThr 1088  
 Db 3425 CTGAGTCAAGCCTGGAAGACAGACAAAGAGGAGCAATGCG---AGCCATGTATATAC 3481  
 QY 1089 LeuAlaLeuIleThrTYrAlaLeuSerSerValGlySer---ProValAlaGlyAla 1107  
 Db 3482 AAAGCATGCTGCTTAATGCTTTTGTCCCTGGCAGGTAAACAGACAAAGAAAGAAATA 3541  
 QY 1108 LeuAsnMetLeuThrTrpArgAlaGluGlnGluGlyGlyMetGlnPheTrp----- 1124  
 Db 3542 CTCAGTCACTTAATAGAGAAAGCTGTGAAGAAAGCACTCTGTCCATTTGGAGGCGCT 3601  
 QY 1125 ValSerSerGluSerLysLeuSerAspSerTrpGlnProArg-----SerLeuAspIle 1142  
 Db 3602 CAGAAACCAAGGACAGAGTGGGCAATTTTAAGAACCCAGAGCTCCCTGTGTAGAGTG 3661  
 QY 1143 GluValAlaAlaTYrAlaLeuLeuSerHisPheLeuGlnPheGln-----ThrSer 1159  
 Db 3662 GAGATGACATCTTAATGCTCTCTGCT---TATCTCAGGCGCCAGCCAGCCCAACTCG 3718  
 QY 1160 GluGlyIlePro-----IleMetArgTrpLeuSerArgGlnArgAsnSerLeu 1175  
 Db 3719 GAGAGCTGACCTCTGCAACCAACATCTGTGAAGTGAACAGACAGACAGAAATCCAG 3778  
 QY 1176 GlyGlyPheAlaSerThrGlnAspThrThrValAlaLeuLysAlaLeuSerGluPheAla 1195  
 Db 3779 GCGGCTTCTCTCCACCAAGAGACAGAGTGTGTGCTCTCATGCTCTGTCCAAATATGGA 3838

```

Qy 1196 AlaLeuMetAnthrGluArgThrAsnIleGlnValThrValThrGlyProSerSerPro 1215
Db 3893 GCCCGCATTATACC---AGGACTGGGAAGGCTGCACAGTACTATCAGTCTTCAAGG 3895
Qy 1216 SerProLeuAlaValValGln----- 1222
Db 3896 ACATTTCACGCAATTCGAAAGTGACAAACAATGCTGTTACTGCACAGGCTCA 3955
Qy 1223 -----PrometAlaValAsnIleSerAlaAsnIlePheGlyPheAlaIleCys 1238
Db 3956 TTGCCAGAGCTGCTCGGGGAATACAGCATGAAGTGAAGAGAGATGTGTACTCTC 4015
Qy 1239 GlnLeuAsnValValTyrAsnValIlySalSerGlySerSerArgArgArgSerIle 1258
Db 4016 CAGACTCTCTGAATACAAATAT-----CTC 4042
Qy 1259 GlnAsnGlnGluAlaPheAspLeuAspValAlaValIlyS-----GluAsnIlyAsp 1275
Db 4043 CCAGAAAGAGAAAGACTCCCTTGTCTTGAAGTGCAGACTGCTGCTCAACTGTGTAT 4102
Qy 1276 AspLeuAsn---HisValAspLeuAsnValCysThrSerPheSerGlyProGly----- 1292
Db 4103 GAACCCAAAGCCACACACAGCTCCAAATCTCCCTAAGTGTACATACAGGAGCCGC 4152
Qy 1293 ---ArgSerGlyMetAlaLeuMetGluValAsnLeuSerGlyPheMetValProSer 1311
Db 4163 TCTGCTCTCAACATGCTGATGTGTGAAGTGTCTGCTGCTTCACTTCCCTGAG 4222
Qy 1312 GluAlaIleSerLeu-----SerGluThrValIlyValGluTyrAspHisGly 1338
Db 4223 CCACAGAGAAATGCTTGAAGAATCTAACCATGAGCCGACAGAAAGTACAGCAGAAC 4282
Qy 1329 LysLeuAsnLeuTyrLeuAspSerValAsnGluThrGlnPheCysValAsnIleProAla 1348
Db 4283 CATGCTGATTTACTCTGATAGGTGTCAATCATGACATGACAGCTGTTCTTCAAGGT 4342
Qy 1349 ValArgAsnPheIlyValSerAsnThrGlnAspAlaSerValSerIleValAspTyr 1368
Db 4343 CTGCAAGATGTCCGATGAAGATCTCAACACAGCATAGTGAAGTCTATGATTACTAC 4402
Qy 1369 GluProArgArgGlnAlaValArgSerTyrAsnSer 1380
Db 4403 GAGACGATGAGTTGCAATGCTGATGATCAATGCT 4438

RESULT 8
US-10-331-496A-19
Sequence 19, Application US/10331496A
Publication No. US20030228305A1
GENERAL INFORMATION:
APPLICANT: FRANTZ, GRETCHEN
APPLICANT: HILLAN, KENNETH J.
APPLICANT: PHILLIPS, HEIDI S.
APPLICANT: POLAKIS, PAUL
APPLICANT: SMITH, VICTORIA
APPLICANT: SPENCER, SUSAN D.
APPLICANT: WILLIAMS, P. MICKEY
APPLICANT: WU, THOMAS D.
APPLICANT: ZHANG, ZEMIN
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DIAGNOSIS AND
FILE REFERENCE: P5014R1-PCT
CURRENT APPLICATION NUMBER: US/10/331,496A
CURRENT FILING DATE: 2002-12-30
PRIOR APPLICATION NUMBER: US 60/345,444
PRIOR FILING DATE: 2002-01-02
PRIOR APPLICATION NUMBER: US 60/351,885
PRIOR FILING DATE: 2002-01-25
PRIOR APPLICATION NUMBER: US 60/360,066
PRIOR FILING DATE: 2002-02-25
PRIOR APPLICATION NUMBER: US 60/362,004
PRIOR FILING DATE: 2002-03-05
PRIOR APPLICATION NUMBER: US 60/366,869
PRIOR FILING DATE: 2002-03-20

```

```

PRIOR APPLICATION NUMBER: US 60/366,284
PRIOR FILING DATE: 2002-03-21
PRIOR APPLICATION NUMBER: US 60/368,679
PRIOR FILING DATE: 2002-03-28
PRIOR APPLICATION NUMBER: US 60/404,809
PRIOR FILING DATE: 2002-08-19
PRIOR APPLICATION NUMBER: US 60/405,645
PRIOR FILING DATE: 2002-08-21
NUMBER OF SEQ ID NOS: 95
SEQ ID NO 19
LENGTH: 4577
TYPE: DNA
ORGANISM: Homo sapien
US-10-331-496A-19

Alignment Scores:
Pred. No.: 3,56e-133 Length: 4577
Score: 1440.00 Matches: 451
Percent Similarity: 46.20% Conservative: 266
Best Local Similarity: 29.06% Mismatches: 551
Query Match: 19.60% Indels: 284
DB: Gaps: 59

US-10-020-095-4 (1-1428) x US-10-331-496A-19 (1-4577)
Qy 12 LeuLeuCysValCysThrAlaAlaLeuAlaValAlaProGlyProArgPheLeuValThr 31
Db 86 CTCCTGCTCTCTCCGCCACAGACGCTGCTCTGGAAGAACGAGATATGTTCTG 145
Qy 32 AlaProGlyIleIleArgProGlyGlyAsnValThrIleGlyValGluLeuLeuGluHis 51
Db 146 GTCCCTCTCTCTCCACT--GAGACCATGAGAAAGGCTGTGCTCTGAGCTAC 202
Qy 52 CysProSerGlnValThrValIlySalGluLeuLeuLysThrAlaSerAsnLeuThrVal 71
Db 203 CTGAATGACAGAGTGTAGTGTCTCTGAGTCTGTCAAGGAAACAGAGCTC 262
Qy 72 ---SerValLeuGluIlyGluGly---ValPheGluIlySerPheLysThrLeuThr 89
Db 263 TTCACACTACTGAGGCGGAGATGACGTCTCACTGCTGCCCTC----- 310
Qy 90 LeuProSerLeuProLeuAsnSerAlaAspGlu-----IleTyrGluLeuAlaVal 106
Db 311 -----GCTGTCCCAAGTCTTATCATGATGAGAGGTATGTTCTCACTGTCAGGTG 364
Qy 107 ThrGlyArgThrGlnAspGluIleLeuPheSerAsnSerThrArgLeuSerPheGluThr 126
Db 365 AAAGACCAACCCAGAA-----TTTAAAGCGGACCAAGATGATGTTAAGAAC 415
Qy 127 LysArgIleSerValPheIleGlnThrAspIlyAlaLeuTyrIlyProGlyAspGluVal 146
Db 416 GAGGACAGTGTGCTTGTTCAGACAGACAAATCATCTACAAACAGGCGAGACGTG 475
Qy 147 LysPheArgIleValThrLeuPheSerAspPheIlyProTyrIlyThrSerLeuAsnIle 166
Db 476 AAATTGCTGTGTCTCATGATGAATAACTTACACCCCTGATGATGATTCACCA 535
Qy 167 Leu---IleLysAspProLysSerAsnLeuIleGlnGluThrPheLeuSerGlnIleAsp 185
Db 536 GTATACCTTCAAGATCCCAAGAAATCGCATGACCAATGAGAGATTCACGTTAGAG 595
Qy 186 LeuGlyValIleSerLysThrPheGlnLeuSerSerHisProIleLeuGlyAspTyrSer 205
Db 596 GGTGGCTCAAGCAATTTCTTTCCCTTATCAGAGCCCTTCCAGGCTCTCTACAG 655
Qy 206 IleGlnValGlnValAsnAsp-----GlnThrTyrTyrGlnSerPheGlnValSerGlu 223
Db 656 GTGGTGTAACAAGAAATCAGGTGAGAGACAGACAC---CCTTTCACCGTGAGGAA 712
Qy 224 TyrValLeuProLysPheGluValThrLeuGlnThrProLeuTyrCysSerMetAsnSer 243
Db 713 TTGTCTCTCCCAAGTTGAAGTGAAGTGAACAGTGCACAAAGATATCACTCTTGAA 772

```



QY 244 LysHisLeuAsnGlyThrIleThrAlaIleuTyThrThrGlyLeuProValIleuSerGlyAsp 263  
 Db 773 GAAGAGATGATATGTATGAGTGTGGCTTATACATATGAGGAGGCTGTCCCTGGACAT 832  
 QY 264 ValIleuThrPheLeuProLeu----- 271  
 Db 833 GTGAGCTGAGCATTTGGCAGAAAAGTATGATGACGCTTCGCACTGCCAGGTGAGATTCA 892  
 QY 272 -----SerPheTyrGlyLeu 276  
 Db 893 CAGGCTTTGTGAGAAATTTCAGTGCACAGCTTAAAGCCATGCGCTCTTATACGCA 952  
 QY 277 LysIleAsnIleThrIleThrPheIleAsnGlySerIleAsnIlePheAsnIle 296  
 Db 953 GTAAAA-----ACCAAGGCTTCCAGCTGAGAGG-----AAGAG 988  
 QY 297 GluGluMetLysAsnValMetAspSerSerAsnGlyLeuSerGluTyrLeuAspLeuSer 316  
 Db 989 TATGAATGAAA-----CTTCACACTGAGGCCAGATCCAA 1024  
 QY 317 SerProGlyProValGluIleLeuThr-----ThrValIleThrGluSerVal 331  
 Db 1025 GAAGAAGCAACACTGTGTGAAATTTGACTGAGAGCACTCCAGTAATTCAGAGAACATA 1084  
 QY 332 ThrGlyIleSerArgAsnValSerThrAsnValPhePheIleGluHisAspTyrIleIle 351  
 Db 1085 ACCAAACTCTCA-----TTGTGAAGTGAAGTCACTTCGACAGGGA-----ATT 1132  
 QY 352 GluPhePheAspTyrThrThrValIleuLysProSerLeuAsnPheThrAlaThrValIle 371  
 Db 1133 CCCTTCTTTGG-----CAG 1147  
 QY 372 ValIleThrArgAlaAspGlyAsnGluLeuThrIleGluGluArgArgAsnValIleIle 391  
 Db 1148 GTGGCGCTAGTATGAGAAAGCGCTCCCTATACCA-----ATAAAGTCATA 1195  
 QY 392 ThrValIleThrGluArgAsnTyrThrGluTyrTyrPheSerIleAsnSerGlyAsnGluLys 411  
 Db 1196 TTCATC---AGAGAAATGATAGCAAACTATTAC-----TCCATGCTCAACGAGTAGAG 1246  
 QY 412 MetGluAlaValGluLysIleAsnTyrThr-----ValIleProGluSerGlyThrPhe 428  
 Db 1247 CATGGCGTTGTACAGTCTCTATCAACACCAACGATATGAGTACTCTCTTACTGTT 1306  
 QY 429 LysIleGluPhePro-----IleLeuGluAspSer 438  
 Db 1307 AAGGATTAATCAAGAGATGTAGTCCCTGTACGCTTACCAAGTGGGTGCAGAAAGACAC 1366  
 QY 439 SerGluLeuGluLeuLysAlaTyrPheLeuGlySerIleSerSerMetAlaValHisSer 458  
 Db 1367 GAAGAGGCACATCACTGCTTAT----- 1390  
 QY 459 LeuPheLysSerProSerIleThrTyrIleGluLeuLysThrArgAspGluAsnIleLys 478  
 Db 1391 CTTGTGTTCTCCCAAGCAAGACTTGTCCACCTTGACCCCATGCTCATGAATCAACC 1450  
 QY 479 ValGlySer-----ProPheGluLeuValValSerGlyAsnLys-----Arg 492  
 Db 1451 TGTGGCCATACTCAAGACTCCAGCAGCATATATTTGAATGAGAGCACTGCTGGGG 1510  
 QY 493 LeuLysGluLeuSer-----TyrMetValIleSerArgGlyGluLeuValAlaValGly 510  
 Db 1511 CTGAAGAACTCTCTTATATATCTGATTAATGCAAAAGAGGCACTTCCGAACTGGG 1570  
 QY 511 -----LysGluAsnSerThr-----MetPheSerLeuThr 520  
 Db 1571 ACTCATGAGACTGCTGTGTGAAGCAGAAAGCATGAGGCCATTTTCCATCAATCCCT 1630  
 QY 521 ProGluAsnSerTyrPheProLysAlaCysValIleValIleTyrTyrIleGluAspArgGly 540  
 Db 1631 GTGAATCAAGCAATTCCTCTGCTGCGGTGCTCATCTATGCTGTATTCCTTACCGGG 1690  
 QY 541 GluIleIleSerAspValLeuLysIleProValGluLeuValPheLysAsnLysIleLys 560

Db 1691 GACGTGATTTGGGATTTGCAAAATATGATGTGAAAATTGTCTGGCCACAAGGTGAT 1750  
 QY 561 LeuTyrTyrPheLysValIleAlaGluProSerGluLysValSerLeuArgIleSerVal 580  
 Db 1751 TTGAGCTTACGCCATCAACAAAGTCTCCAGCTTCAACAGGCCACCTGAGTGCACAGC 1810  
 QY 581 ThrGluProAspSerIleValGlyIleValAlaValAspLysSerValAsnLeuMetAsn 600  
 Db 1811 GCT---CCTCACTCCGTCTGGCGCCCTCCGTGTGTGACCAAAAGCTGTCTCATGAAG 1867  
 QY 601 AlaSerAsnAspIleThrMetGluAsnValValHisGluLeu---GluLeuTyrAsnThr 619  
 Db 1868 CCGTAGTGCAGCTCTCGCGCTCTGATTACAACTGTATCAAGAAAGAAAGACTCTCACT 1927  
 QY 620 GlyTyr-----TyrLeuGlyMet 625  
 Db 1928 GGCCTCCCTGGCGCTTGAATGACAGAGCAGATGAGACTGCATCATGCTATATGTC 1987  
 QY 626 PheMetAsnSer-----PheAla 631  
 Db 1988 TATATTAATGGAATCACTATATCTCACTATCACTACAAATGAAAAGATATGTACGC 2047  
 QY 632 ValPheGluGluCysGlyLeuTyrValLeuThrAspAlaAsnLeuThrLysAspTyrIle 651  
 Db 2048 TTCCTAGAGACATGCGCTTAAAGCATTCACCAACTCAAAAGATTGTAACCCAAATG 2107  
 QY 652 AspGlyValIleTyrAspAsnAlaGlu-----Tyr 660  
 Db 2108 TGTCCACAGCTTCAACAGTATGAATGATGATGAGCACTGAGGTATGAGTTTAT 2167  
 QY 661 AlaGluArgPheMetGluGluAsnGluGlyHisIleValAspIleHisAspPheSerLeu 680  
 Db 2168 GAGTCAAGATGTAAATGGAAGAGGCCATGACAGCTCGGTGATGTTGAAGAG----- 2218  
 QY 681 GlySerSerProHis-----ValArgLysHisAspProGluThrTyrIleThrLeu 697  
 Db 2219 -----CCTCACAGGAGACGCTTGAAGAGTACTTCCCTGAGACATGATCTGGAT 2269  
 QY 698 AspThrAsnMetGlyTyrArgGlyLeuGluGluPheGluValThrValProAspSerIle 717  
 Db 2270 TTGGTGTGTGAATCACTGACAGGGGTGCTGAGTAGAGTAAACAGTCCCTGACACATC 2329  
 QY 718 ThrSerTyrValAlaIleThrGlyPheValIleSerGluAspLeuGlyLeuThrThr 737  
 Db 2330 ACCGATGGAAGGCGGGGCTTCTGCTGTGAAGAGCTGGAATTGATCTCTTCC 2389  
 QY 738 ThrProValGluLeuGluAlaPheGluProPhePheIlePheLeuAsnLeuProTyrSer 757  
 Db 2390 ACT---GCTCTCTCCAGGCTTCCAGGCTTCTTGTGAGAGCTTACAAATGCTTACTCT 2446  
 QY 758 ValIleArgGlyGluGluPheAlaLeuGluIleThrIlePheAsnTyrLeuLysAspAla 777  
 Db 2447 GTGATTCGTGAGAGGCTTCACTCAAGCCCAAGGCTTCAAACTTCCCAATATGC 2506  
 QY 778 ThrGluValLysValIleIleGluLysSerAspLysPheAspIleLeuMetThrSerSer 797  
 Db 2507 ATCCGGGTCAAGTGTCACTGAAAGCTTCCCGCTTCTCTGCTGCCAGTGAAGAG 2566  
 QY 798 Glu-----IleAsnAlaThrGlyHisGluInThrLeuValProSer 812  
 Db 2567 GAACAAGGCGCTCACTGATCTGTCCAAACGGGCGCA----- 2605  
 QY 813 GluAspGlyAlaThrValLeuPheProIleArgProThrHisLeuGlyGluIleProIle 832  
 Db 2606 -----ACTGTCTCTGGGCAAGTAAACCCCAAGTCAATTAAGAAATGCAATTTC 2653  
 QY 833 ThrValThrAlaLeuSer-----ProThrAla 841  
 Db 2654 ACTGTGACGCGAGAGGCACTGAGTCTTAAGAGCTGTGTGGAATGAGAGTGTCTTCAATT 2713  
 QY 842 -----SerAspAlaValThrGluMetIleLeuValIleAlaGluGlyIle 856

Db 2714 CCTGAACAGGAGAAAGACACAGCTCATCAAGCTCTGTTGGTGAACCTGAAGACTA 2773  
 Qy GluYserSerThrSerGlnSerIleLeuLeuAspLeuThrAspAsnArgLeuGlnSerThr 876  
 Db 2774 GAGAAAGAAACACACTTCAACTCCCTACTTGTTCATCAAGTGGGAGGTTCT- 2827  
 Qy LeuYThrLeuSerPheSerPheProProAsnThrValThrGlySerGlnArgValGln 896  
 Db 2828 --GAAAGATTATTCCTGAAACTGCCACCAATGCTGAGAAAGATTCGCCGAGCTTCT 2884  
 Qy IleThrAlaIleGlyAspValLeuGlyProSerIleAsnGlyLeuAlaSerLeuIleArg 916  
 Db 2885 GCTCAAGTTTGGGACATATATAGCTCTCTGCCATGCAAAAACACAAATCTTCTCCAG 2944  
 Qy MetProTyrGlyCysGlyGlnGlnAsnMetIleAsnPheAlaProAsnIleTyrIleLeu 936  
 Db 2945 ATGCCCTTAGCTGGTGGAGACAGAAATGTGCTCTTGTCTCTTACATCTATGTAAGT 3004  
 Qy AspTyrLeuThrLysLysGlnLeuThrAspAsnLeuLysGlnValAlaLeuSerPhe 956  
 Db 3005 GATTATCTAAATGAAACACAGAGCTTACTCCAGAGCTCAAGTCCAAAGCCATTGGCTAT 3064  
 Qy MetArgGlnGlyTyrGlnArgGlnLeuLeuTyrGlnArgGlnArgGlySerPheSerAla 976  
 Db 3065 CTCAACACTGGTTCACAGACAGCTTGAACCTAACAACACTATGAGCTCTTACAGCAC 3124  
 Qy PheGly--AsnTyrAspProSer--GlySerThrTyrLeuSerAlaPheValLeuArg 994  
 Db 3125 TTTGGGAGGCATATGACAGAACACAGGACCAACCTGGCTCAGAGCTTGTCTGTAAG 3184  
 Qy CysPheLeuGlnAlaAspProTyrIleAspIleAsnGlnAsnValLeuIleArgThrTyr 1014  
 Db 3185 ACTTTGGCCAAAGCTCGAGCCTACATCTTCATGCATGAAAGACACATTAACCAAGCCCTC 3244  
 Qy ThrTyrLeuLysGlyHisGlnLysSerAsnGlyLysPheTyrProGlyArgValIle 1034  
 Db 3245 ATATAGCTCTCCAGAGGACAGAAAGCAATGCTGTTCAGAGCTCTGGCTCACTGCTC 3304  
 Qy HisSerGlnLeuGlnGlyGlyAsnLysSerProValThrLeuThrAlaTyrIleValThr 1054  
 Db 3305 AACAAATGCCATTAAGAGAGAGAGTAGAAGATGAGACCTTCGCTTATATACCATC 3364  
 Qy SerLeuLeuGlyTyrArgLysTyrGlnProAsnIleAsnLysSerIleHisPhe 1074  
 Db 3365 GCCCTTGGAGATTCCTCAAGTCACTCACTGCTGTGTCGGAATGCCCTGTTTTC 3424  
 Qy LeuGlnSer-----GluPheSerArgGlyIleSerAsnTyrThr 1088  
 Db 3425 CTGAGTCAAGCTCGGAGAGACACACAAAGAGGAGCCATGTC--AGCCATGTATATACC 3481  
 Qy LeuAlaLeuIleThrTyrAlaLeuSerSerValGlySer--ProLysAlaLysGlnAla 1107  
 Db 3482 AAAGCACTGCTGGCTATGCTTTGCTCTGGCAGGTATACAGACAAAGAGAGAGAGTA 3541  
 Qy LeuAsnMetLeuThrTyrArgAlaGlnGlnGlnGlyLysMetGlnPheTyr----- 1124  
 Db 3542 CTCAGTCACTTAATGAGAAAGCTGTAAAGAAACAACCTGTCATTTGGAGGCGCT 3601  
 Qy ValSerSerGlnSerLysLeuSerAspSerTyrGlnProArg-----SerLeuAspIle 1142  
 Db 3602 CAGAAACCCAGACCCAGCTGGGAGCATTTTACGAACCCAGAGCTCCCTCTGAGAGTGG 3661  
 Qy GluValAlaAlaTyrAlaLeuLeuSerHisPheLeuGlnPheGln-----ThrSer 1159  
 Db 3662 GAGATGACATCTTATGCTCTCTCTGCT--TATCTCAAGCGCCACAGCCAGCCCACTCG 3718  
 Qy IleMetArgTyrLeuSerArgGlnArgAsnSerLeu 1175  
 Db 3719 GAGGACCTGACCTTCGCAACCAACATCTGTGAAGTGCATCAGAGACAGCAAAATGCCAG 3778  
 Qy GlyGlyPheAlaSerThrGlnAspThrThrValAlaLeuLysValAlaLeuSerGlnPheAla 1195  
 Db 3779 GCGCGTTTCTCTCTCAACCCAGACACAGTGTGTGCTCTCATGCTCTGCAAAATATGGA 3838

Qy 1196 AlaLeuMetAsnThrGluArgThrAsnIleGlnValThrValThrGlyProSerSerPro 1215  
 Db 3839 GCGGCAACATTTACC--AGAGCTGGAGAGCTGCACAGGTGACTATCCAGTCTTACAGG 3895  
 Qy 1216 SerProLeuAlaValValGln----- 1222  
 Db 3896 ACAATTTCCAGAAATTCGAAATGACACACAAATCGCTGTACTGCAGAGCTTCA 3955  
 Qy 1223 -----PrometAlaValAsnIleSerAlaAsnGlyPheGlyPheAlaIleCys 1238  
 Db 3956 TTGGCAGAGCTGCTGGGAAATACAGCATGAAGTGCAGAGAGAGATGTGCTACTTC 4015  
 Qy 1239 GlnLeuAsnValValTyrAsnValLysAlaSerGlySerSerArgArgArgSerIle 1258  
 Db 4016 CAGACCTCTGTAATACATAT-----CTC 4042  
 Qy 1259 GlnAsnGlnLysAlaPheAspLeuAspValAlaValLys-----GluAsnLysAsp 1275  
 Db 4043 CCAAGAAAGAGAGATTCCTTGTGCTTGAAGTGCAGAGCTGCTGCTCAAACTTGAT 4102  
 Qy 1276 AspLeuAsn--HisValAspLeuAsnValCysThrSerPheSerGlyProGly----- 1292  
 Db 4103 GAACCCAAAGCCACACACAGCTTCAAATCTCCCTTAAGTGTCACTTACACAGAGAGCTCC 4162  
 Qy 1293 ---ArgSerGlyMetAlaLeuMetGluValAsnLeuSerGlyPheMetValProSer 1311  
 Db 4163 TCGGCTCCCAACATGGGAGATCGTGAATGTGAATGTGCTCTGCTCATCTCCCGTAAG 4222  
 Qy 1312 GluAlaIleSerLeu-----SerGlnThrValLysValGlnTyrAspHisGly 1328  
 Db 4223 CCAACAGTGAATATGCTTGAAGATCTAACATGATGAGCGGACAGACAGACAGCAC 4282  
 Qy 1329 LysLeuAsnLeuTyrLeuAspSerValAsnGlnThrGlnPheCysValAsnIleProAla 1348  
 Db 4283 CATGCTTGATTTTACCTTGATTAAGTGTCAATCAAGACTGAGCTTGTCTTACAGGTT 4342  
 Qy 1349 ValArgAsnPheLysValSerAsnThrGlnAspAlaSerValSerIleValAspTyr 1368  
 Db 4343 CTCGAGATGCTCCAGTAAGATGATCAACCAAGCCATGATGAAGTGTATGATTTACTAC 4402  
 Qy 1369 GluProArgArgGlnAlaValArgSerTyrAsnSer 1380  
 Db 4403 GAGACGATGATTTGCAATCGCTGAGTACAAATGCT 4438

RESULT 9  
 US-09-873-319-408  
 ; Sequence 408, Application US/09873319A  
 ; Publication No. US20030134324A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Mungert, William E.  
 ; APPLICANT: Kulikarni, Prakash  
 ; APPLICANT: Getzenberg, Robert H.  
 ; APPLICANT: Waga, Iwao  
 ; TITLE OF INVENTION: Identifying Drugs for and Diagnosis of Benign Prostatic  
 ; TITILE OF INVENTION: Hyperplasia Using Gene Expression Profiles  
 ; FILE REFERENCE: 44921-5029-US  
 ; CURRENT APPLICATION NUMBER: US/09/873,319A  
 ; EARLIER FILING DATE: 2001-06-05  
 ; EARLIER APPLICATION NUMBER: US 60/223,323  
 ; NUMBER OF SEQ ID NOS: 755  
 ; SOFTWARE: PatentIn Ver. 2.1  
 ; SEQ ID NO 408  
 ; LENGTH: 4577  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 ; FEATURE:  
 ; OTHER INFORMATION: Genbank Accession No. US20030134324A1 M11313  
 US-09-873-319-408  
 Alignment Scores:

Pred. No.: 3,566-133 Length: 4577  
 Score: 1440.00 Matches: 451  
 Percent Similarity: 46.20% Conservative: 266  
 Best Local Similarity: 29.06% Mismatches: 551  
 Query Match: 19,604 Indels: 284  
 DB: 13 Gaps: 59

US-10-020-095-4 (1-1428) x US-09-873-319-408 (1-4577)

QY 12 LeuLeuCyValCyThrAlaLeuAlaValAlaProGlyProArgPheLeuValThr 31  
 |||||  
 |||||  
 DB 86 CTCTTGTCCTCTCTGCGCCACAGACGCTCATGTCCTGGAACCGAGATATGTTCTG 145  
 QY 32 AlaProGlyIleIleArgProGlyGlyAsnValThrIleGlyValGluLeuLeuGlnHis 51  
 |||||  
 |||||  
 DB 146 GTCCCTCCCTGCTCCACACT--GAGACCACTGAGAGAGGCGTGTCTCTGAGCTAC 202  
 QY 52 CySProSerGlnValThrValValAlaGluLeuLeuThrAlaSerAsnLeuThrVal 71  
 |||||  
 |||||  
 DB 203 CTGAATGAGACAGTGAAGTGAAGTCTCTGAGCTGTCTGAGGGGAAACAGAGCTTC 262  
 QY 72 ---SerValLeuGlnAlaGluGly---ValPheGluLeuGlySerPheLeuThr 89  
 |||||  
 |||||  
 DB 263 TTCACCTGACCTGAGCGGAGAGATGACGTAATCTCCACTGTGCGCTTC----- 310  
 QY 90 LeuProSerLeuProLeuAsnSerAlaAspGlu-----IleTyGluLeuArgVal 106  
 |||||  
 |||||  
 DB 311 -----GCTGTCCCAAGCTTCATCTCATGAGAGAGTATGTTCTCACTGTCCAGT 364  
 QY 107 ThrGlyArgThrGlnAspGluIleLeuPheSerAsnSerThrArgLeuSerPheGluThr 126  
 |||||  
 |||||  
 DB 365 AAGAGCCAAACCCAGAA-----TTTAAAGACCGAGACAGATGATGTTAGTAAAC 415  
 QY 127 LysArgIleSerValPheIleGlnThrAspLysAlaLeuThrLysProGlnGluVal 146  
 |||||  
 |||||  
 DB 416 GAGGACAGTGTGTCCTTGTCCAGACAGCAATCATCTACAAACAGGAGGACAGCTG 475  
 QY 147 LysPheArgIleValThrLeuPheSerAspPheLysProTyLysThrSerLeuAsnIle 166  
 |||||  
 |||||  
 DB 476 AATTTTCGTGTGTCTCCATGATGATAAACTTCCACCCCTGATGATGATGATTCACAT 535  
 QY 167 Leu---IleLysAspProLysSerAsnLeuIleGlnIleThrLeuSerGlnIleSerAsp 185  
 |||||  
 |||||  
 DB 536 GATATCATTCAGAGATCCCAAGGAAATCCCATGCGACATGAGAGAGTTTCCAGTAA 595  
 QY 186 LeuGlyValIleSerLysThrPheGlnLeuSerSerHisProIleLeuGlyAspTrpSer 205  
 |||||  
 |||||  
 DB 596 GGTGGCTTCAGCAATTTTCTTTTCCCTCATCATCAGAGCCCTTCCAGGGCTTCCACAG 655  
 QY 206 IleGlnValGlnValAsnAsp-----GlnThrTyLysThrIleSerPheGlnValSerGlu 223  
 |||||  
 |||||  
 DB 656 GTGGGTGTCACAGAAATCAGGTGAGAGACAGAGAC--CTTTCACTCGTGAAGAA 712  
 QY 224 TyrValLeuProLysPheGlnValThrLeuGlnThrProLeuTyLysSerMetAsnSer 243  
 |||||  
 |||||  
 DB 713 TTTGTTCTTCCCAAGTTTGAAGTACAGTAACAGTGCAGAAATATATCATCATCTTGAA 772  
 QY 244 LysHisLeuAsnGlyThrIleThrAlaLysTyThrTyGlyLysProValLysGlyAsp 263  
 |||||  
 |||||  
 DB 773 GAAGAGATGAATGTATCAGTGTGTGCTTATACATATGAGAGAGCTGTCCCTGAGCAT 832  
 QY 264 ValThrLeuThrPheLeuProLeu----- 271  
 |||||  
 |||||  
 DB 833 GTGACTGTGAGCATTTTGCAAGAAATGATGATGAGCTTCCGACTGCCAGGTGAAGATTCA 892  
 QY 272 -----SerPheTrpGlyLys 276  
 |||||  
 |||||  
 DB 893 CAGGCTTTCTGTGAGAAATTCAGTGACAGCTAAACAGCATAGCGCTTATATACGAA 952  
 QY 277 LysLysAsnIleThrLysThrPheLysIleAsnGlySerAlaAsnPheSerPheAsnAsp 296  
 |||||  
 |||||  
 DB 953 GTTAAAT-----ACCAAGGTCTTCCAGCTGAAGAG-----AAGGAG 988

QY 297 GluGluMetLysAsnValMetAspSerSerAsnGlyLeuSerGluTyLysAspLeuSer 316  
 |||||  
 |||||  
 DB 989 TATGAATGATAA-----CTTCACATGAGGCCAGATTCOA 1024  
 QY 317 SerProGlyProValGluIleLeuThr-----ThrValThrGluSerVal 331  
 |||||  
 |||||  
 DB 1025 GAAGAGAAACAGTGTGGAATGTAGCTGAGAGGAGCTCCAGTGAATATCAAGAAACATA 1084  
 QY 332 ThrGlyIleSerArgAsnValSerThrAsnValPhePheLysGlnHisAspTyIleIle 351  
 |||||  
 |||||  
 DB 1085 ACCAACTCTCA---TTTGTGAAGTGAAGCTCATCTTCCAGAGGA-----ATT 1132  
 QY 352 GluPhePheAspTyThrThrValLeuLysProSerLeuAsnPheThrAlaThrValLys 371  
 |||||  
 |||||  
 DB 1133 CCTTCTTTGGG-----CAG 1147  
 QY 372 ValThrArgAlaAspGlyAsnGlnLeuThrLeuGlnGluArgArgAsnValValIle 391  
 |||||  
 |||||  
 DB 1148 GTGCCCTAGTAGAGAGGAAAGCGCTCCCTATACCA-----AATTAAGTCATA 1195  
 QY 392 ThrValThrGlnArgAsnTyThrGluTyTrpSerGlySerAsnSerGlyAsnGlnLys 411  
 |||||  
 |||||  
 DB 1196 TTCATC---AGAGAAATGAAAGCAAACTATTAC-----TCCATGTCTACCGAGATGAG 1246  
 QY 412 MetGluAlaValGln---LysIleAsnTyThr-----ValProGlnSerGlyThrPhe 428  
 |||||  
 |||||  
 DB 1247 CATGGCTTTGTATCATTTCTATACACCAACCAAGTTATGGTACCTCTTATCTGT 1306  
 QY 429 LysIleGluPhePro-----IleLeuGluAspSer 438  
 |||||  
 |||||  
 DB 1307 AGGGTCAATTACAAAGATGATGCTCTGTACGGCTACCAAGTGGGTGCAGAAACAC 1366  
 QY 439 SerGluLeuGlnLeuLysAlaTyThrPheLeuGlySerLysSerMetAlaValHisSer 458  
 |||||  
 |||||  
 DB 1367 GAAGAGCATCATCATCTTAT----- 1390  
 QY 459 LeuPheLysSerProSerLysThrTyIleGlnLeuLysThrArgAspGluAsnIleLys 478  
 |||||  
 |||||  
 DB 1391 CTGTGTCTTCCCAAGCAAGAGCTTGTCCACTTGAGCCCATGTCTCATGAATCACTACC 1450  
 QY 479 ValGlySer-----ProPheGluLeuValValSerGlyAsnLys-----Arg 492  
 |||||  
 |||||  
 DB 1451 TGTGGCCATTAACAACAGATGATGCTCTGTACGGCTACCAAGTGGGTGCAGAAACAC 1510  
 QY 493 LeuLysGluLeuSer-----TyrMetValValSerArgGlyGlnLeuValAlaValGly 510  
 |||||  
 |||||  
 DB 1511 CTGAAGAGCTCTCTTATTATATCTGATATGCGAAAGGAGGAGCATTTGCGAACTGGG 1570  
 QY 511 -----LysGlnAsnSerThr-----MetPheSerLeuThr 520  
 |||||  
 |||||  
 DB 1571 ACTCATGACCTCTGTGAAGCAGAGACATGAAGGCGCATTTTCCATCTCAATCCCT 1630  
 QY 521 ProGluAsnSerTrpThrProLysAlaLysValIleValTyTyIleGluAspAspGly 540  
 |||||  
 |||||  
 DB 1631 GTGAAGTCAAGATGCTCTGTGCGCTGCGTGTGCATCTATGCTTTTACTACCGGG 1690  
 QY 541 GluIleIleSerAspValLeuLysIleProValGlnLeuValPheLysAsnLysIleLys 560  
 |||||  
 |||||  
 DB 1691 GACGTGATGGGATTCGCAAAATATGATGTGTAATAATGTCTGCGCAACAGAGTGAT 1750  
 QY 561 LeuTyTrpSerLysValLysAlaGluProSerGluLysValSerLeuArgIleSerVal 580  
 |||||  
 |||||  
 DB 1751 TTGACCTTCAAGCCATCAACAAAGTCTCCAGGCTTACACGCCACCTGAGATCAGCGG 1810  
 QY 581 ThrGlnProAspSerIleValGlyIleValAlaValAspLysSerValAsnLeuMetAsn 600  
 |||||  
 |||||  
 DB 1811 GCT---CCATGCTCGTGGGCGCTCGGTGTGTGAGCAAAAGCTGTCTGTATGAAG 1867  
 QY 601 AlaSerAsnAspIleThrMetGluAsnValValHisGluLeu---GluLeuTyTrpAsnThr 619  
 |||||  
 |||||  
 DB 1868 CCTGATGCTGAGCTCTCGGCGCTCGGTTTACAACTGTACACAGAAAGAGACCTTCACT 1927  
 QY 620 GlyTyR-----TyrLeuGlyMet 625

Db 1928 GGCCTCCCTGGGCTTGAATGACCGAGCAGTGAAGTGCATCAATCGTCAATGATGC 1987  
 Qy 626 PheMetAnser-----PheAla 631  
 Db 1988 TATTTAATGGAATCACAATATCTCCAGTATCAAGTCAAAATGAAAGATATGTACAGC 2047  
 Qy 632 ValPheGlnGlySerGlyLeuTrpValLeuThrAspAlaAsnLeuThrIleAspTrpIle 651  
 Db 2048 TTCTAGAGGACATGGGCTTAAGAGCATTCACCAATCAAAAGATTCTGTAACCAAAATG 2107  
 Qy 652 AspGlyValIleAspAsnAlaGln-----Trp 660  
 Db 2108 TGTCCACAGCTTCAACAGTATGAATGACATGACCTGAAGGTCTACGTGATTTTAT 2167  
 Qy 661 AlaGlnArgPheMetGlnGluAsnGlnGlyIleValAspIleHisAspPheSerLeu 680  
 Db 2168 GAGTCAGATGTAATGGGAAGAGCCCATGACCGCTGGTGATGTGAAG----- 2218  
 Qy 681 GlySerSerProHis-----ValArgIleHisPheProGlnTrpIleTrpLeu 697  
 Db 2219 -----CCTCACAGGAGAACCGTACGAAAGTACTTCCCTGAGACATGATCTGGGAT 2269  
 Qy 698 AspThrAsnMetGlyIleArgIleIleTrpGlnGluPheGluValIleProAspSerIle 717  
 Db 2270 TTGGTGGTGTAACTACAGCAGGGGTGGCTGAGTAGAGTAAACAGTCCCTGACACCATC 2329  
 Qy 718 ThrSerTrpValAlaThrGlyPheValIleSerGlnPheLeuGlyLeuThr 737  
 Db 2330 ACCAGTGAAGAGGAGGGGCTTCTGCTGTGAAGATGCTGACTGTGATCTCTTC 2389  
 Qy 738 ThrProValGlnLeuGlnAlaPheGlnProPhePheIlePheLeuAsnLeuProTrpSer 757  
 Db 2390 ACT---GCTCTCTCCAGCCTTCCAGCCTTCTTGTGAGACTTACAACTGCTTACTCT 2446  
 Qy 758 ValIleArgGlyGlnGluPheAlaLeuGluIleThrIlePheAsnTrpLeuAspAla 777  
 Db 2447 GTGATTCGTGGAGAGGCTTACACTACAGCCAGCGCTCAACTACCTTCCCAAAATGC 2506  
 Qy 778 ThrGluValIleValIleGlnIleGlySerAspIleAspIleLeuMetThrSer 797  
 Db 2507 ATCCGGTTCAGTGTGAGCTGGAAGCCCTCTCCGCTCTCTGCTCCAGTGGAGAG 2566  
 Qy 798 Gln-----IleAsnAlaThrGlyHisGlnGlnThrLeuValProSer 812  
 Db 2567 GAACAACGCTCTCATCTGCTGTGCAACGGGGCGCA----- 2605  
 Qy 813 GluAspGlyAlaThrValLeuPheProIleArgProThrHisIleGlnIleProIle 832  
 Db 2606 -----ACTGTGCTCTGGGAGTAACTCCAAAGTCAATTAGAAATGTGAATTC 2653  
 Qy 833 ThrValThrAlaLeuSer-----ProThrAla 841  
 Db 2654 ACTGTGAGCGAGGAGCTAGAGTCTCAAGAGCTGTGGAGCTGAGGTGCTTCAATT 2713  
 Qy 842 -----SerAspAlaValThrGlnMetIleLeuValIleValIleGlnIle 856  
 Db 2714 CTTGAACACGGAAGAAAGACACAGTCAAGCTCTGTTGGTGAACCTGAAGGACTA 2773  
 Qy 857 GlnIleSerTrpSerGlnSerIleLeuLeuAspLeuThrAspAsnArgIleGlnSerThr 876  
 Db 2774 GAGAAAGAAACATCACTCCCTACTTGTCTCATCAGGTGTGAGTTCT----- 2827  
 Qy 877 LeuIleThrLeuSerPheSerPheProProAsnThrValIleThrGlySerGlnArgValGln 896  
 Db 2828 ---GAAACAATTATCTCGAAACTGCCACCAATGTGTAGAGAAATCTGCCGCGCTTCT 2884  
 Qy 897 IleThrAlaIleGlyAspValLeuGlyProSerIleAsnGlyLeuAlaSerLeuIleArg 916  
 Db 2885 GTCTCAGTTTGGAGACATTTAGGCTTGCACATGCAAAACACACAAATCTTCTCAG 2944  
 Qy 917 MetProTrpGlyCysGlyGlnGlnAsnMetIleAsnPheAlaProAsnIleTrpIleLeu 936  
 |||||  
 |||||

Db 2945 ATGCCCTATGCTGTGGAGAGCAAAATATGCTCTTGTGCTTCAATCTATGATGC 3004  
 Qy 937 AspTrpLeuThrIleValIleGlnLeuThrAspAsnLeuGlyValIleLeuSerPhe 956  
 Db 3005 GATTATCTAATGAATAACACAGCAGCTTATCTCCAGAGTCAAGTCCAAAGGCATTTGGCTAT 3064  
 Qy 957 MetArgGlnIleTrpGlnArgGlnLeuLeuTrpGlnArgGlnAspIleSerPheSerAla 976  
 Db 3065 CTCACACCTGTTATCCAGAGACAGTTGAACCTAACAACTATATGATGCTCTACAGACAC 3124  
 Qy 977 PheGly---AsnTrpAspProSer---GlySerThrTrpLeuSerAlaPheValIleuArg 994  
 Db 3125 TTGGGAGAGCAATATGACAGAAACAGGCAACCTGCTCAACCTTGTGTCTGAAG 3184  
 Qy 995 CysPheLeuGlnAlaAspProTrpIleAspIleAspGlnAsnValLeuHisArgTrpTrp 1014  
 Db 3185 ACTTTTGCCCAAGCTGAGGCTTACATCTTATGATGAAAGACACATTCACCAAGCCCTC 3244  
 Qy 1015 ThrTrpLeuGlySerHisGlnIleSerAsnGlyGlnPheTrpAspProGlnArgValIle 1034  
 Db 3245 ATATGGCTCTCCAGAGGAGAGAGCAATGCTGTTTCAGAGCTCTGGGTCACTGCTC 3304  
 Qy 1035 HisSerGlnLeuGlnGlyGlyAsnIleSerProValIleLeuThrAlaTrpIleValIle 1054  
 Db 3305 AACAAATGCCATAAAGGAGAGTGAAGATGAAGTACCTCTCCGCTATATCACCATC 3364  
 Qy 1055 SerLeuLeuGlyTrpArgIleTrpArgIleTrpGlnProAsnIleAspValGlnIleSerIleHisPhe 1074  
 Db 3365 GCCCTTGTGAGATTCCTCTCACATGCTACCTCTGTTGTCCGAATGCCCTGTTTGC 3424  
 Qy 1075 LeuGlnSer-----GluPheSerArgGlyIleSerAspAsnTrpTrp 1088  
 Db 3425 CTGAGTCAAGCCTCGAAGACAGCAAGAAAGGAGCAATAGCC---AGCATGTATATAC 3481  
 Qy 1089 LeuAlaLeuIleThrTrpAlaLeuSerSerValGlySer---ProIleAlaValGlnAla 1107  
 Db 3482 AAAGCACTGTGCTCATGTGCTTGTGCTGCGCAGGTACCAAGCAAGAGGAAGTA 3541  
 Qy 1108 LeuAsnMetLeuThrTrpArgIleGlnGlnGlnGlyIleMetGlnPheTrp 1124  
 Db 3542 CTCAAGTCACTTATAGAGAGCTGTGAAGAAAGACAACTCTCTCCATTTGGAGCCCT 3601  
 Qy 1125 ValSerSerGlnSerIleSerAspSerTrpGlnProArg-----SerLeuAspIle 1142  
 Db 3602 CAGAAACCAAGGACACAGAGGGGATTTTACAGAACCCAGGCTCCCTGTGAGGTG 3661  
 Qy 1143 GluValAlaAlaTrpAlaLeuLeuSerHisPheLeuGlnPheGln-----ThrSer 1159  
 Db 3662 GAGATGACATCTCATGTGCTCTCTGCT---TATCTCACGCGCCAGCCAGCCCAACCTCG 3718  
 Qy 1160 GlnGlyIlePro-----IleMetArgTrpLeuSerArgGlnArgAsnSerLeu 1175  
 Db 3719 GAGGACCTGACCTCTGCAACCAATCGTGAATGATGATCAAGAGCAAGAAATGCGCAG 3778  
 Qy 1176 GlyIlePheAlaSerThrGlnAspThrTrpValAlaLeuValAsnLeuSerGlnPheAla 1195  
 Db 3779 GGGGTTTCTCCCTCCACCAAGACACAGAGTGGCTCTCAATGCTCTGTGCAAAATGGA 3838  
 Qy 1196 AlaLeuMetAsnThrGlnArgThrAsnIleGlnValIleValIleThrGlyProSerSerPro 1215  
 Db 3839 GCCGCCACATTTTAC---AGGACTGGGAAGGCTGACAGAGTACTATCCAGTCTTCAGGG 3895  
 Qy 1216 SerProLeuAlaValValGln----- 1222  
 Db 3896 ACAATTTCCAGCAAAATTCACAGTGAACAACAAATGCGCTTACTGACAGAGTCTCA 3955  
 Qy 1223 -----ProMetAlaValAsnIleSerAlaAsnGlyPheGlyPheAlaIleCys 1238  
 Db 3956 TTGCCAGAGCTGCTGGGAATACAGCATGAAGTACAGAGAAAGATGTCTACCTC 4015  
 Qy 1239 GlnLeuAsnValValIleTrpAsnValIleValAsnGlySerSerArgArgArgSerIle 1258  
 |||||  
 |||||

QY 1259 GlnAsnGlnAlaPheAspValAlaVallys-----GlnAsnLysAsp 1275  
 Db 4043 CCAGAAAGAGAGAGTCCCTTTCAGAGTCCAGACTGCTGCTCAAACTTGAT 4102  
 QY 1276 AspleAsn---HisAlaAspleAsnValCysThrSerPheSerGlyProGly----- 1292  
 Db 4103 GAACCCAAAGCCACCCAGCTTCCAAATCTCCCTTAAGTCTACAGAGACCCG 4162  
 QY 1293 ---ArgSerGlyMetAlaLeuMetGluValAsnLeuSerGlyPheMetValProser 1311  
 Db 4163 TCTGCTCCACATGGCCGATCGTGAATGGAAGATGCTCTGAGCTTCTCCCTGAAG 4222  
 QY 1312 GluAlaLeuSerLeu-----SerGluThrValLysValGluTyraPheIsgly 1328  
 Db 4223 CCACAGTAAATGCTTGAAGATCTAACATGTGAGCCGACAGACAGCAGAAC 4282  
 QY 1329 LysLeuAsnLeuTyrlleuAspSerValAsnGluThrGlnPheCysValAsnIleProAla 1348  
 Db 4283 CATGCTTATTTACCTTGATTAAGGTGCAATTCAGACATGAGCTTCTTCACGGTT 4342  
 QY 1349 ValArgAsnPheLysValSerAsnThrGlnAspAlaSerValSerIleValAspTyrr 1368  
 Db 4343 CTGCAAGATGCTCCAGTAAAGATCTCAACCAAGCATATGTGAAGTCTATTAATCTAC 4402  
 QY 1369 GluProArgArgGlnAlaValArgSerTyraAsnSer 1380  
 Db 4403 GAGACGATGAGTTTGCATTCGCTGAGTCAATTCCT 4438  
 RESULT 10  
 US-09-960-706-654  
 / Sequence 654, Application US/09960706  
 / Publication No. US20030134280A1  
 / GENERAL INFORMATION:  
 / APPLICANT: Mungen, William E.  
 / TITLE OF INVENTION: Identifying Drugs for and Diagnosis of Benign Prostatic Hyperplasia  
 / FILE REFERENCE: 44921-5029-01US  
 / CURRENT APPLICATION NUMBER: US/09/960,706  
 / PRIOR FILING DATE: 2001-09-24  
 / PRIOR APPLICATION NUMBER: 60/223,323  
 / PRIOR FILING DATE: 2000-08-07  
 / PRIOR APPLICATION NUMBER: 09/873,319  
 / NUMBER OF SEQ ID NOS: 1124  
 / SOFTWARE: PatentIn Ver. 2.1  
 / SEQ ID NO 654  
 / LENGTH: 4577  
 / TYPE: DNA  
 / ORGANISM: Homo sapiens  
 / FEATURE:  
 / OTHER INFORMATION: Genbank Accession No. US20030134280A1 M11313  
 US-09-960-706-654  
 Alignment Scores:  
 Pred. No.: 3,566-133 Length: 4577  
 Score: 1440.00 Matches: 451  
 Percent Similarity: 46.20% Conservative: 266  
 Best Local Similarity: 29.06% Mismatches: 551  
 Query Match: 19.60% Indels: 284  
 DB: 13 Gaps: 59  
 US-10-020-095-4 (1-1428) x US-09-960-706-654 (1-4577)  
 QY 12 LeuLeuCysValCysThrAlaAlaLeuAlaValAlaProGlyProArgPheLeuValThr 31  
 Db 86 CTCCTGGTCTCTCTGCGCACAGACGCTCAGTCTCGAAGAACCGAGATATGTTCTG 145  
 QY 32 AlaProGlyIleIleAlaArgProGlyGlyAsnValThrIleGlyValGluLeuLeuGluHis 51  
 Db 146 GTCCCTCCCTCTCTCACACT--GAGACCACTGAGAGAGGGCTGTCTCTTGAGCTAC 202  
 QY 52 CysProSerGlnValThrValLysAlaGluLeuLeuTyrlleuValSerAsnLeuThrVal 71

Db 203 CTGAATGAGACAGTCACTGTAACTGCTTCTGAGTCTGTCAAGGAGAAACAGAGCCTC 262  
 QY 72 ---SerValLeuGlnAlaGluGly---ValPheGluLysGlySerPheLysThrLeuThr 89  
 Db 263 TTCATGACCTGAGAGCGGAGATGACGATCTCCACTGTGTGCCCTTC----- 310  
 QY 90 LeuProSerLeuProLeuAsnSerAlaAspGlu-----IleTyrlGluLeuArgVal 106  
 Db 311 -----GCTGTCCTCAAGCTTCACTTCAATGAGAGAGTAAATTTCTCTCACTGCTCAAGT 364  
 QY 107 ThrGlyArgThrGlnAspGlnIleLeuPheSerAsnSerThrArgLysSerPheGluThr 126  
 Db 365 AAAGACCAACCCCAAGAA-----TTAAGAAGCGGACACAGTGTGTTAAGAAC 415  
 QY 127 LysArgIleSerValPheIleGlnThrAspLysAlaLeuTyrlleuProGlyGlnGluVal 146  
 Db 416 GAGACAGCTGTGCTTTGTCCAGACAGACAAATCTACAAACCAAGGACAGACAGT 475  
 QY 147 LysPheArgIleValThrLeuPheSerAspPheLysProTyrlleuThrSerLeuAsnIle 166  
 Db 476 AAATTTCTGTTCTCTCCATGATGAAATCTTTCACCCCTCAATGATGATTCACACTA 535  
 QY 167 Leu---IleLysAspProLysSerAsnLeuIleGlnGlnThrLeuSerGlnGlnSerAsp 185  
 Db 536 GTATCATTCAGATCCCAAGAGAAATCGCATGCGACATGSCAGAGTTTCCAGTTAAG 595  
 QY 186 LeuGlyValIleSerLysThrPheGlnLeuSerSerHisProIleLeuGlyAspTrpSer 205  
 Db 596 GGTGGCTTCACAAATTTTCTTCTTCCCTCATCAGAGCCCTTCAGAGGCTCCATCAAG 655  
 QY 206 IlegValAlaGlnValAsnAsp-----GlnThrTyrlleuGlnSerPheGlnValSerGlu 223  
 Db 656 GTGGGTGTCACAGAAATTCAGTGAAGAGACAGACGAC--CCTTTACCTGAGAGAA 712  
 QY 224 TyrValLeuProLysPheGlnValThrLeuGlnThrProLeuTyrlleuSerMetAsnSer 243  
 Db 713 TTTGTTCTTCCCAATTTGAAGTACAGATTAACAGGCCAAAGATATATACACTCTTGGA 772  
 QY 244 LysHisLeuAsnGlyThrIleThrAlaLysTyrlleuTyrlleuGlyLysProValLysGlyAsp 263  
 Db 773 GAAGAGATGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 832  
 QY 264 ValThrLeuThrPheLeuProLeu----- 271  
 Db 833 GTGACTGTGAGCATTTGCAAGAAATGATAGTACGCTTCCAGCTGCCACGTCAGATTC 892  
 QY 272 -----SerPheTrpGlyLys 276  
 Db 893 CAGGCTTCTGAGAAATTCAGTGAAGACGTAACACCCATGCTGCTTCTATAGCA 952  
 QY 277 LysLysAsnIleThrLysThrPheLysIleAsnGlySerAlaAsnPheSerPheAsnAsp 296  
 Db 953 GTAAAA-----ACCAAGTCTTCCAGCTCAAGAG-----AAGAG 988  
 QY 297 GluGluMetLysAsnValMetAspSerSerAsnGlyLeuSerGluTyrlleuAspLeuSer 316  
 Db 989 TATGAATGAAGAA-----CTTCACTAGAGGCCCAATTC 1024  
 QY 317 SerProGlyProValGluIleLeuThr-----ThrValThrGluSerVal 331  
 Db 1025 GAAGAGAGACAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1084  
 QY 332 ThrGlyIleSerArgAsnValSerThrAsnValPhePheLysGlnHisAspTyrlleu 351  
 Db 1085 ACCAAACTCTCA--TTGTGAAGTGAATCAGTCACTTTCAGACAGGA-----ATT 1132  
 QY 352 GluPhePheAspTyrlleuThrValLeuLysProSerLeuAsnPheThrAlaThrValLys 371  
 Db 1133 CCTTCTTGGG-----CAG 1147  
 QY 372 ValThrArgAlaAspGlyAsnGlnLeuThrLeuGluGluArgArgAsnAsnValValIle 391

Db 1148 GTGGCCCTAGATGAGAAAGCGCTCCATACCA-----AATAAGCATATA 1195  
Qy 1392 ThrValThrGlnArgAsnThrTrpGluTrpSerGlySerAsnSerGlyAsnGlnLys 411  
Db 1196 TTATC---AGAGAAATGAAGCAAACTATTAC-----TCCATGCTACCAACGAGAG 1246  
Qy 412 MetGlnAlaValGln---LysIleAsnThr-----ValProGlnSerGlyThrPhe 428  
Db 1247 CATGCCCTTGACAGTTCTCTATCAACACCAACGTTATGGTACCTCTTACTGTT 1306  
Qy 429 LysIleGluPhePro-----LysGluAspSer 438  
Db 1307 AGGGTCATTACAAAGATCGTAGTCCCTGTTACGGCTACCAAGGCTTCAGAAAGAAC 1366  
Qy 439 SerGluLeuGlnLeuLysAlaThrPheLeuGlySerSerSerMetAlaValHisSer 458  
Db 1367 GAAGAGGACATCACACTGCTTAT----- 1390  
Qy 459 LeuPheLysSerProSerLysThrTrpIleGlnLeuLysThrArgAspGlnAsnIleLys 478  
Db 1391 CTGTGTTCTCCCAAGCAAGAGCTTGTCCACCTTGACCCATGCTCATGAATACCC 1450  
Qy 479 ValGlySer-----ProPheGluLeuValLysSerGlyAsnLys-----Arg 492  
Db 1451 TGTGGCCATCTCAGACAGTCACAGGCACATTATATTCGAATGAGAGCCCTCGGGG 1510  
Qy 493 LeuLysGluLeuSer-----TyrMetValLysSerArgLysGlnLeuValAlaValGly 510  
Db 1511 CTGAAGAGCTCTCTTTTATTATCTGATATGCGAAGGAGGACATTGTCCGAACCTGG 1570  
Qy 511 -----LysGlnAsnSerThr-----MetPheSerLeuThr 520  
Db 1571 ACTCATGACTGCTGTGACAGAGAGAACATGAAGGCCATTTTTCATCTCATTCCT 1630  
Qy 521 ProGluAsnSerTrpThrProLysAlaLysValIleValTyrTrpIleGlnAspAspGly 540  
Db 1631 GTGAGATGACATTTGCTCTCGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1690  
Qy 541 GluIleIleSerAspValLeuLysIleProValGlnLeuValPheLysAsnLysIleLys 560  
Db 1691 GACGTGATGGGAGATTCTGCAAAATATGATGTTGAAATTTGTGCGCAACAGGTGAT 1750  
Qy 561 LeuTrpSerLysValLysAlaGluProSerGluLysValSerLeuArgIleSerVal 580  
Db 1751 TTGAGCTTCACGCCATCACAAGTCTCCAGCTCCACAGCCACCTGCGATCAGACGG 1810  
Qy 581 ThrGlnProAspSerIleValGlyIleValAlaValAspLysSerValAsnLeuMetAsn 600  
Db 1811 GCT---CCTCAGTCCGTCTGCGGCTCCGTCTGTCGACCAAGGCTGCTCATGAG 1867  
Qy 601 AlaSerAsnAspIleThrMetGluAsnValAlaHisGluLeu---GluLeuTyrAsnThr 619  
Db 1868 CTTGATGCTGAGCTCTCGGCGCTCTCGGTTTACACCTGCTACCAAGAAAGACCTCTCT 1927  
Qy 620 GlyTyr-----TyrLeuGlyMet 625  
Db 1928 GCGTTCCTCGGCGCTTGAATGACACGAGACGATGAAGCTGCATCAATCGTCAATATTC 1987  
Qy 626 PheMetAsnSer-----PheAla 631  
Db 1988 TATATTATGCAATCATATATCTCACTACATCAATCAATGAAGAAAGATATGACGC 2047  
Qy 632 ValPheGlnGluLysGlyLeuTrpValLeuThrAspAlaAsnLeuThrLysAspTyrIle 651  
Db 2048 TTCTTAGAGGACATGGGCTTAAAGGCAATTCACCAATCAAGATTCTGTAACCCCAAAATG 2107  
Qy 652 AspGlyValTyrAspAsnAlaGlu-----Tyr 660  
Db 2108 TGTCCACAGCTTCACACATATGAATATGATGACCTGCAAGGCTTACGTTAGCTTTTAT 2167  
Qy 661 AlaGluArgPheMetGluGlnAsnGluLysIleValAspIleHisAspPheSerLeu 680  
Db 2168 GAGTCAGATGTAATGGAAGAGGCAATGACGCTGTCATGTTGAAGAG----- 2218

Qy 681 GlySerSerProHis-----ValArgLysHisPheProGluThrTrpIleTrpLeu 697  
Db 2219 -----CCTCACAGGAGACCGGATACGAAAGATCTCCCTGACATGATCTGGAT 2269  
Qy 698 AspThrAsnMetGlyTyrArgIleTyrGlnGluPheGluValThrValProAspSerIle 717  
Db 2270 TTGGTGTGTAACTCAGACAGGGGTGGCTGAGTGTAGAGTAAACAGTCCCTGACACATC 2329  
Qy 718 ThrSerTrpValAlaThrGlyPheValIleSerGluAspLeuGlyLeuThrThr 737  
Db 2330 ACCGAGTGAAGGACAGGGGCTTCTGCTCTCTGGAAGTCTGACCTGATCTCTTCC 2389  
Qy 738 ThrProValGluLeuGlnAlaPheGlnProPhePheIlePheLeuAsnLeuProTyrSer 757  
Db 2390 ACT---GCTCTCTCCGAGGCTTCAGACCTCTCTTGTGAGGCTTCAATAGCTTACTCT 2446  
Qy 758 ValIleArgGlyGluGluPheAlaLeuGluIleThrIlePheAsnTyrLeuLysAspAla 777  
Db 2447 GTGATTCGTGAGAGGCTTCACATCAGCCACGCTCTTAACTAATCTTCCCAATGTC 2506  
Qy 778 ThrGluValLysValIleIleGluLysSerAspLysPheAspIleLeuMetThrSerSer 797  
Db 2507 ATCCGGGTGAGTGTGACGTGAGAGCTCTCCGCTCTTCTGCTGCCAGTGGAGAG 2566  
Qy 798 Glu-----IleAsnAlaThrGlyHisGlnGlnThrLeuLeuValProSer 812  
Db 2567 GAACAAAGGCTCAGTCACTGTGCAAGAGGCGGCA----- 2605  
Qy 813 GluAspGlyAlaThrValLeuPheProIleArgProThrHisLeuGlyGluIleProIle 832  
Db 2606 -----ACTGTCCTCGGCGAGTAAACCCAAAGTCAATGAAATGTAATTC 2653  
Qy 833 ThrValThrAlaLeuSer-----ProThrAla 841  
Db 2654 ACTGTAGCCCAAGGACCTAGAGTCAAGCTGTGAGGACGTGAGCTTCACTT 2713  
Qy 842 -----SerAspAlaValThrGlnMetIleLeuValLysAlaGluGlyIle 856  
Db 2714 CTGGAACAGGAGAAAGAACACAGTCAACAGCTCTGTTGTTGAACCTGAAGAGCTA 2773  
Qy 857 GluLysSerTyrSerGlnSerIleLeuLeuAspLeuThrAspAsnArgLeuGlnSerThr 876  
Db 2774 GAGAGGAAACCAATCAATCACTCCCTTGTTCACAGGTGAGGCTTCTT----- 2827  
Qy 877 LeuLysThrLeuSerPheSerPheProAsnThrValThrGlySerGluArgValGln 896  
Db 2828 ---GAAATATATCCCTGAATCTGCCCAAAATGTGTAGAGATGTGCCGAGCTTCT 2884  
Qy 897 IleThrAlaIleGlyAspValLeuGlyProSerIleAsnGlyLeuAlaSerLeuIleArg 916  
Db 2885 GTCTCACTTTTGGAGAGCATTTATAGCTCTGCCATGCAAAACACAAATCTTCTCAG 2944  
Qy 917 MetProTyrGlyCysGlyGlnGlnAsnMetIleAsnPheAlaProAsnIleTyrIleLeu 936  
Db 2945 ATGCCATAGGCTGTGAGAGACAAATATGTCTCTTCTCTCAATCATATGTACTG 3004  
Qy 937 AspTyrIleuThrLysLysGlnLeuThrAspAsnLeuLysGluLysAlaLeuSerPhe 956  
Db 3005 GATTATTAATGAAMAACAGAGCTTACTCAGAGGTCAAGTCAAGGCGCATTTGCTAT 3064  
Qy 957 MetArgGlnGlyTyrGlnArgGluLeuLeuTyrGlnArgLysAspGlySerPheSerAla 976  
Db 3065 CTCAACACTGTTTACAGACAGCTTGAATCACTCAAAACATATATGCTCTCTACAGCAC 3124  
Qy 977 PheGly---AsnTyrAspProSer---GlySerThrTrpLeuSerAlaPheValLeuArg 994  
Db 3125 TTGGGAGAGGATATGACAGAAACAGACAGCTGCTGCTTGTCTGAG 3184  
Qy 995 CysPheLeuGlnAlaAspProTyrIleAspIleAspGlnAsnValLeuHisArgThrTyr 1014  
Db 3185 ACTTTGCCAAGCTCGAGGCTTACATCTTATCATGATGAAGACACACATTACCAAGCCCTC 3244



QY 1015 ThrTrpLeuLysGlyHisGlnLysSerAsnGlyValPheTrpAspProGlyArgValIle 1034  
 DB 3245 AATAGCTCTCCAGAGGAGGAGCAATGCTGTTCCAGAGCTCTGCTGCTACCTGC 3304  
 QY 1035 HisSerGluLeuGlnGlyLysAsnLysSerProValThrLeuThrAlaTyrIleValThr 1054  
 DB 3305 AACATGCCATTAAGGAGAGAGTAGAAGATGAAGTAACTCTCCCTCATATACCACTC 3364  
 QY 1055 SerLeuLeuGlyTyrArgLysTyrGlnProAsnIleAspValGlnLysSerIleHisPhe 1074  
 DB 3365 GCCCTTCTGGAGATTCTCTCCACAGTCACCTCACCTGTTGCGCAATGCCCTGTTTGC 3424  
 QY 1075 LeuGluSer-----GluPheSerArgLysIleSerAspAntyTyr 1088  
 DB 3425 CTGGAGCTCAGCTCGAAGACAGACCAAGAGGAGCAATAGGC---AGCATGTATATACC 3481  
 QY 1089 LeuAlaLeuIleThrTyrAlaLeuSerSerValGlySer---ProLysAlaLysGluAla 1107  
 DB 3482 AAGACAGCTGCTGCTATGCTTCTTGGCTGCGAGGTAAACAGAGCAAGAGAGAAAGTA 3541  
 QY 1108 LeuAsnMetLeuThrTyrArgAlaGlnGlnGlyLysMetGlnPheTrp----- 1124  
 DB 3542 CTCATGCTCTTAATAGAGAGCTGTGAAGAAAGACAATCTCTCCATTGGAGGCCCT 3601  
 QY 1125 ValSerSerGluSerLysSerAspSerTyrGlnProArg-----SerLeuAspIle 1142  
 DB 3602 CAGAAACCAAGGACAGCAGTGGGAGCATTTTACGAACCCAGAGCTCCCTGCTGAGTGG 3661  
 QY 1143 GluValAlaAlaTyrAlaLeuSerHisPheLeuGlnPheGln-----ThrSer 1159  
 DB 3662 GAGATGACATCTATGCTGCTCTGCT---TATCTCAGCGCCAGCCAGCCCAACCTCG 3718  
 QY 1160 GluGlyIlePro-----IleMetArgTrpLeuSerArgGlnArgAsnSerLeu 1175  
 DB 3719 GAGGACTGACCTCTCCACCAACCAATCGTAAGTGCATCAGAGAGAGAGATGCCAG 3778  
 QY 1176 GlyGlyPheAlaSerThrGlnAspThrThrValAlaLeuLysAlaLeuSerGluPheAla 1195  
 DB 3779 GGGGGTTCCTCCACCCAGAGCAGAGTGGTCTCTCATGCTCTGTCCTCAATATGGA 3838  
 QY 1196 AlaLeuMetAsnThrGluArgThrAsnIleGlnValThrValThrGlyProSerSerPro 1215  
 DB 3839 GCGGCACATTTTACC---AGGACTGGAGAGGCTGCAAGGTGATCTATCCAGTCTTCAGG 3895  
 QY 1216 SerProLeuAlaValValGln----- 1222  
 DB 3896 ACATTTTCCAGCAATTCAGAGTGACAAACAAATACGCTGACTGACAGAGTCTCA 3955  
 QY 1223 -----PrometAlaValAsnIleSerAlaAsnGlyPheGlyPheAlaIleCys 1238  
 DB 3956 TTGCCAGAGCTGCTGGGGAATACAGCATGAAGTGAAGAGAGATGTGTCTACCTC 4015  
 QY 1239 GlnLeuAsnValValTyrAsnValLysAlaSerGlySerSerArgArgArgSerIle 1258  
 DB 4016 CAGACCTCTTGAATAATACATAT-----CTC 4042  
 QY 1259 GlnAsnGlnGluAlaPheAspLeuAspValAlaValLys-----GluAsnLysAsp 1275  
 DB 4043 CCAAGAAAGAGAGAGTTCCTTGTGAGAGTCAAGTCACTGCTCAAACTGTGAT 4102  
 QY 1276 AspLeuAsn---HisValAspLeuAsnValCysThrSerPheSerGlyProGly----- 1292  
 DB 4103 GAACCAAAAGCCACACCAAGCTTCCAAATCTCCCTAAGTGTCAACACAGGAGACCC 4162  
 QY 1293 ---ArgSerGlyMetAlaLeuMetGluValAsnLeuSerGlyPheMetValProSer 1311  
 DB 4163 TCTGCTCCAAACATGCGATGCTGTGATGTAAGATGTCTCTGCTTCAATCCCTGAAG 4222  
 QY 1312 GluAlaIleSerLeu-----SerGluThrValLysValLysValGluTyrAspHisGly 1328  
 DB 4223 CCAACAGTGAATAATGCTTGAAGATCAACATGTAGCCGACAGACAGTCAAGCAAC 4282  
 QY 1329 LysLeuAsnLeuTyrLeuAspSerValAsnGluThrGlnPheCysValAsnIleProAla 1348

DB 4283 CATGCTTGATTACTTGATGATAGGTGTCAAAATCAGACACTGAGCTTGTCTTACCGTT 4342  
 QY 1349 ValArgAsnPheLysValSerAsnThrGlnAspAlaSerValSerIleValAspTyr 1368  
 DB 4343 CTGCAGATGTCTCCAGTATAGAGATCTCAACCAAGCATGTATGAATATATATAC 4402  
 QY 1369 GluProArgArgGlnAlaValArgSerTyrAsnSer 1380  
 DB 4403 GAGACGAGATGAGTTTGCAATGCTGATGATCAATGCT 4438  
 RESULT 11  
 US-10-292-081A-3  
 ; Sequence 3, Application US/10292081A  
 ; Publication No. US20030162202A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Kenneth David Becker  
 ; APPLICANT: Genul Veliclelebi  
 ; APPLICANT: Xin Wang  
 ; APPLICANT: Randolph E. Tanzi  
 ; APPLICANT: Lars Berttram  
 ; APPLICANT: Aleister J. Saunders  
 ; TITLE OF INVENTION: SINGLE NUCLEOTIDE POLYMORPHISMS AND MUTATIONS ON ALPHA-2-MACROG  
 ; FILE REFERENCE: 37481-3323  
 ; CURRENT APPLICATION NUMBER: US/10/292,081A  
 ; PRIOR FILING DATE: 2002-11-08  
 ; PRIOR APPLICATION NUMBER: 60/337434  
 ; NUMBER OF SEQ ID NOS: 15  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 3  
 ; LENGTH: 4577  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 US-10-292-081A-3  
 Alignment Scores:  
 Pred. No.: 3,566-133 Length: 4577  
 Score: 1440.00 Matches: 451  
 Percent Similarity: 46.20% Conservative: 266  
 Best Local Similarity: 29.06% Mismatches: 551  
 Query Match: 19.60% Indels: 284  
 DB: 13 Gaps: 59  
 US-10-020-095-4 (1-1428) x US-10-292-081A-3 (1-4577)  
 QY 12 LeuLeuCysValCysThrAlaAlaLeuAlaValAlaProGlyProArgPheLeuValThr 31  
 DB 86 CTCTTGATCTCTCTCCACAGAGCCCTCAGTCTCTGGAAGAAACCGAGATATGTTCTG 145  
 QY 32 AlaProGlyIleIleArgProGlyLysValThrIleGlyValGlnLeuGlnHis 51  
 DB 146 GTCCCTCCCTGCTCCACT---GAGACCAGTGAAGAGGCTGTGCTCTTGTGAGCTAC 202  
 QY 52 CysProSerGlnValThrValLysAlaGluLeuLysThrAlaSerAsnLeuThrVal 71  
 DB 203 CTGAATGACAGACGTATCTATAGTCTCTTGGAGTCTGTCAAGGAAACAGAGCTC 262  
 QY 72 ---SerValLeuGlnAlaGlnGly---ValPheGluLysGlySerPheLysThrLeuThr 89  
 DB 263 TTCACTGACCTGAGAGCGGAGATAGCATCTCCACTGTGTGCGCTTC----- 310  
 QY 90 LeuProSerLeuProLeuAsnSerAlaAspGlu-----IleTyrGlnLeuArgVal 106  
 DB 311 -----GCTGTCCAAAGCTTTCATCAATGAGAGGTAAATGTTCTCACTGCAAGT 364  
 QY 107 ThrGlyArgThrGlnAspGluLeuPheSerAsnSerThrArgLeuSerPheGluThr 126  
 DB 365 AAGAGCAACCAAGAA---TTTAAAGACGCGACCAACAGTATGTATTAAAGAAC 415  
 QY 127 LysArgIleSerValPheIleGlnThrAspLysValAlaLeuTyrLysProLysGlnGluVal 146  
 DB 416 GAGGACAGTCTGTCTTGTCTGACAGACAAATCAATCTCAAAACCAAGGCGACAGTGG 475

147 LysPheArgIleValThrIleuPheSerAspPheLysProTyrIleThrSerIleuAsnIle 166  
 476 AAATTCCTGCTGTCTCATGATGATAAACTTTCACCCCTGATGATGATTCACCTA 535  
 167 Leu--IleLysAspProLysSerAsnIleuIleGlnIleuPheSerGlnIleuSerAsp 185  
 536 GTATTCATTCAGATCCCAAGAAATGCGATCCGACAAATGCGAGAGTTCCAGATTAGAG 595  
 186 LeuGlyValIleSerIleuThrPheGlnIleuSerSerHisProIleuGlyAspTyrSer 205  
 596 GGTGCTCCCAAGCAATTTCTTTCTTCTTCCTCATACAGAGCTTCCAGAGCTCCACAG 655  
 206 IleGlnIleGlnValAsnAsp-----GlnThrTyrTyrGlnSerPheGlnValSerGlu 223  
 656 GTGCGTGCACGAATAATCAGTGCAGAGACAGACAGC-----CCTTACACCTGAGAGAA 712  
 224 TyrValIleuProLysPheGluValThrIleuGlnIleuProLeuTyrCysSerMetAsnSer 243  
 713 TTTGTTCTTCCCAAGTTTGAAGTACAGATAACAGTCCAAAGATATACATCATCTTGGAA 772  
 244 LysHisIleuAsnGlyThrIleThrValIleuTyrTyrIleuProValIleGlyAsp 263  
 773 GAAGAGATGAATGTATCAGTGTGGCCTATACATATGAGAGCTGTCCCTGGACAT 832  
 264 ValThrIleuThrPheLeuProLeu----- 271  
 833 GTGACTGTGACATTTGCGAAGAAATAGTAGACCTTCCGACCTGCGACGTAAGATTCA 892  
 272 -----SerPheTyrGlyLys 276  
 893 CAGGCTTCTGTGAAATTCAGTGCAGCTAAACAGCCATGCGTCTTCTATACAGCA 952  
 277 LysLysAsnIleThrIleuThrPheLysIleAsnGlySerAlaAsnPheSerPheAsnAsp 296  
 953 GTAAAA-----ACCAGAGCTCTTCCAGCTGAAGAG-----AAGAG 988  
 297 GluIleuMetLysAsnValMetAspSerSerAsnGlyLeuSerGluTyrLeuAspLeuSer 316  
 989 TATGAAATGAAA-----CTTCACATGAGAGCCCAAGATCAA 1024  
 317 SerProGlyProValIleuIleuLeuThr-----ThrValThrGluSerVal 331  
 1025 GAAGAGAGAACAGTGTGGAATTGACTCGAAGAGCAGTCCAGTGAATACAAAGAACATA 1084  
 332 ThrGlyIleSerAsnValSerThrAsnValPhePheLysGlnHisAspTyrIleIle 351  
 1085 ACCAAACTCTCA--TTTGTGAAGTGACTCACACTTTCGACAGGGA-----ATT 1132  
 352 GluPhePheAspTyrThrThrValIleuLysProSerIleuAsnPheThrAlaThrValLys 371  
 1133 CCCTTCTTTGGG-----CAG 1147  
 372 ValThrArgAlaAspGlyAsnGlnIleuThrIleuGlnIleuArgAsnAsnValIle 391  
 1148 GTGGCCTAGTATAGATGGAAGGCGTCCCTATACAA-----AATAAAGTCTATA 1195  
 392 ThrValThrGlnAsnValThrIleuTyrTyrSerGlySerAsnSerGlyAsnGlnLys 411  
 1196 TTTATC--AGAGGAATGAAGCAAACTATTAC-----TCCAAATGCTACACAGGATCAG 1246  
 412 MetGlnAlaValGln--LysIleAsnTyrThr-----ValProGlnSerGlyThrPhe 428  
 1247 CATGCCCTGTATACATTTCTATACACACACCAACGTTATGGGATCTCTTATCTTAT 1306  
 429 LysIleGluPhePro-----IleLeuGluAspSer 438  
 1307 AGGGTCATTTACAGAGATGATGCTCCGTACGCTACAGGCTACAGTGGGTCTCAAGAAACAC 1366  
 439 SerGluLeuGlnIleuLysAlaTyrPheLeuGlySerLysSerMetAlaValHisSer 458  
 1367 GAAGAGGACATCACTACCTTAT----- 1390

459 LeuPheLysSerProSerIleuThrIleGlnIleuLysThrArgAsnIleLys 478  
 1391 CTGTGTCTTCCCCAACCAAGAGCTTGTCCACTGACCCATGTCTCATGAACATCACCC 1450  
 479 ValGlySer-----ProPheGluLeuValValSerGlyAsnLys-----Arg 492  
 1451 TGTGGCCATCTCAGACAGTCCAGGACATTAATATCTGAATGAGGACCCCTGCTGGG 1510  
 493 LeuLysGluLeuSer-----TyrMetValIleSerArgIleGlnIleuValAlaValGly 510  
 1511 CTGAAGAGCTCTCTTATTTATCTATATGACCAAGGAGCATTTGCCAATCGG 1570  
 511 -----LysGlnAsnSerThr-----MetPheSerLeuThr 520  
 1571 ACTCATGAGCTGCTTGAAAGACAGAAAGACATGAAGGCCATTTTTCATCTCATCCCT 1630  
 521 ProGluAsnSerThrProLysValIleCysValIleValIleTyrIleGluAspArgGly 540  
 1631 GTGAAGTCACACATTCCTCTGCTGCTGCTGCTGCTCATATGCTGTTTACCTACCGG 1690  
 541 GluIleIleSerAspValIleuLysIleProValIleIleuValPheLysAsnLysIleLys 560  
 1691 GACGTGATTTGGGATTTCTGCACAAATATGATGATTTGAAATTTGCTGCGCAACAGTGGAT 1750  
 1751 TTGAGCTTACGCCCATCACAAGATCTCCAGCTCACACCCCATCGGACATCACAGCG 1810  
 581 ThrGlnProAspSerIleValIleGlyIleValAlaValAspLysSerValAsnLeuMetAsn 600  
 1811 GCT---CCTCAGTCCGCTCGCGCCCTCGCTGCTGAGCAAAAGCGTCTCATATGAG 1867  
 601 AlaserAsnAspIleThrMetGluAsnValValHisGluLeu---GluLeuTyrAsnThr 619  
 1868 CCTGATCTGAGCTCTGCGGCTCTGCTTACACCTGCTACACAGAAAGACCTTACT 1927  
 620 GlyTyr-----TyrLeuGlyMet 625  
 1928 GGCCTCCCTGGGCTTGAATGACACAGAGCATGAAGATCATCATCTCATTAATGTC 1987  
 626 PheMetAsnSer-----PheAla 631  
 1988 TATATATATGAAATCATCATTAATCTCAGTATCAAGTACAAATGAAGATATGATACAGC 2047  
 632 ValPheGlnIleCysGlyLeuTyrValIleuThrAspAlaAsnLeuThrLysAspTyrIle 651  
 2048 TTCTTAGAGACATGGGCTTTAAAGCATTCACCAACTCAAAAGATTCTTAACCCAAATG 2107  
 652 AspGlyValTyrAspAlaIleGlu-----Tyr 660  
 2108 TGTCCACAGCTTCAACAGTATGAATGCATGACCTGAAGTCTAGCTGATGTTTAT 2167  
 661 AlaGluArgPheMetGluIleuAsnGlyHisIleValAspIleHisAspPheSerLeu 680  
 2168 GAGTCAGATTAATGGAAGAGGCGATGACGCTGCTGATGTGAAGG----- 2218  
 681 GlySerSerProHis-----ValArgLysHisPheProGluThrTyrIleTyrLeu 697  
 2219 -----CCTCACAGGAGACCGGCTTGAAGAGTCTCTGAGATGATGATTTGGAT 2269  
 698 AspThrAsnMetGlyTyrArgIleTyrGlnIleuPheGluValThrValProAspSerIle 717  
 2270 TTGGTGTGTGTAACATCAGAGAGGAGGCTGAGAGTGAAGATTAACAGTCCCTGACACATC 2329  
 718 ThrSerThrValAlaThrGlyPheValIleSerGluAspLeuGlyLeuGlyLeuThr 737  
 2330 ACCAGTGAAGGACAGGCGCTTCTGCTGCTGCTGCTGATGATCTGATCTCTTCC 2389  
 738 ThrProValGluLeuGlnAlaPheGlnProPhePheIlePheLeuAsnLeuProTyrSer 757  
 2390 ACT---GCCCTCTCCAGGCTTCCAGCCCTTCTTGTGAGACTTACAAAGCCTTACTCT 2446  
 758 ValIleArgIleGluIleuPheAlaLeuGluIleThrIlePheAsnTyrLeuLysAspAla 777



APPLICANT: Xin Wang  
 APPLICANT: Randolph E. Tanzi  
 APPLICANT: Lars Bertam  
 APPLICANT: Aleister J. Saunders  
 TITLE OF INVENTION: SINGLE NUCLEOTIDE POLYMORPHISMS AND MUTATIONS ON ALPHA-2-MACROGLOBULIN  
 FILE REFERENCE: 37481-3323  
 CURRENT APPLICATION NUMBER: US/10/292,081A  
 CURRENT FILING DATE: 2002-11-08  
 PRIOR APPLICATION NUMBER: 60/337434  
 PRIOR FILING DATE: 2001-11-09  
 NUMBER OF SEQ ID NOS: 15  
 SOFTWARE: FastSeq for Windows Version 4.0  
 SEQ ID NO: 5  
 LENGTH: 4577  
 TYPE: DNA  
 ORGANISM: Homo sapiens  
 US-10-292-081A-5

Alignment Scores:  
 Pred. No.: 3,566-133 Length: 4577  
 Score: 1440.00 Matches: 451  
 Percent Similarity: 46.20% Conservative: 261  
 Best Local Similarity: 29.06% Mismatches: 551  
 Query Match: 19.60% Indels: 284  
 DB: 13 Gaps: 59

US-10-020-095-4 (1-1428) x US-10-292-081A-5 (1-4577)

QY	12	LeuIeuCyValCyThThrAlaIaLeuAlaValAlaProGlyProArgPheLeuValThr	31
DB	86	CTCTTGTCCTCTCTCCAGACAGCGCTCAGTCTTGAAAAAGCAGTAATAGTGTCTG	145
QY	32	AlaProGlyIleIleArgProGlyIaValThrIleGlyValGluLeuLeuGluHis	51
DB	146	GMCCCTCCGCTCCACACT--GAGACCACTGAAGAGGCGTGTCTTCTGAGCTAC	202
QY	52	CysProSerGlnValThrValIleAlaGluLeuLeuIleThrAlaSerAsnLeuThrVal	71
DB	203	CTGAATGAGACAGTCACTGAAGTCTTCTTGAGTGTGTCAGGGGAAACAGAGCTCT	262
QY	72	--SerValLeuGluAlaGluGly--ValPheGluIuGlySerPheLeuThr	89
DB	263	TTCACTGACCTGGAGCGGAGATGACGTACTCCACTGTGCGCTTC-----	310
QY	90	LeuProSerLeuProLeuAsnSerAlaAspGlu-----IleTyrGluLeuArgVal	106
DB	311	-----GCTGCCCAAACTCTTCATCCATGAGAGCAATGATTCCTCACTGCCAAGTG	364
QY	107	ThrGlyArgThrGlnAspGluIleLeuPheSerAsnSerThrArgLeuSerPheGluThr	126
DB	365	AAAGGACCAACCCAGAA-----TTTAAAGACCGGACACAGTATGATTAAACAC	415
QY	127	LyAsnGlyIleSerValPheIleGlnThrAspIysAlaLeuTyrIleProGlyGlnGluVal	146
DB	416	GAGACAGACTGTGCTTTGTCCAGACAGACAATCAATCTCAACACGAGGACAGACTG	475
QY	147	LysPheArgIleValThrLeuPheSerAspPheIleProTyrIleThrSerLeuAsnIle	166
DB	476	AAATTTCGTGTGTCTCCATGATGATGAAAACTTTCACCCCTGATGATGATTAATTCACCTA	535
QY	167	Leu--IleIysAspProIysSerAsnLeuIleGlnGlnIleProLeuSerGlnGlnSerAsp	185
DB	536	GTATTCATTCAAGATCCCAAGAAATCGCATCGCAATGCGACAGAGTTTCCAGTTGAG	595
QY	186	LeuGlyValIleSerIleThrPheGlnLeuSerSerHisProIleLeuGlyAspTyrSer	205
DB	596	GGTGGCCCTCAAGCAATTTCTTTTCCCTTCATCAGAGCCCTTCCAGGGGCTCTCAACAG	655
QY	206	IleGlnValGlnValAsnAsp-----GlnThrTyrIleGlnSerPheGlnValSerGlu	223
DB	656	GTGTGTGTACAGAAATACAGGTGGAAGACAGAC---CCTTTACCGGTGAGGAA	712
QY	224	TyrValLeuProIysPheGluValThrLeuGlnThrProLeuTyrCysSerMetAsnSer	243

Db 713 TTGTGTTTCCCAAGTTTGAAAGTACAAAGTACAGTGCACAAAGATATACACATCTTGGA 772  
::: |||  
::: |||  
Oy 244 LysHisLeuAsnGlyThrIleThrAlaLeuSerThrTyrlGlyLysProValLysGlyAsp 263  
::: |||  
::: |||  
Db 773 GAAGAGATGAATGATACAGTGTGTGGCTATACCATATGGGAAGCCTGTCCCTGGACAT 832  
264 ValThrLeuThrPheLeuProLeu----- 271  
|||  
|||  
Db 833 GTGACGTGAGCATTTGCGAAGAAAGTATAGTACGCTTCGACCTGCCAGGTGAAGATTCA 892  
272 -----SerPheTrpGlyLys 276  
|||  
|||  
Db 893 CAGGCTTCTGTGAGAAATTCAGTGCACAGCTAAACGCCATGGCTCTCTATACGAA 952  
277 LysLysAsnIleThrLysThrPheLysIleAsnGlySerAlaAsnPheSerPheAsnAsp 296  
|||  
|||  
Db 953 GTAAAG-----ACCAAGGTCTTCCAGCTGAAGAG-----AAGAG 988  
297 GluGluMetCysAsnValMetAspSerSerAsnGlyLeuSerGluTyrlLeuAspLeuSer 316  
|||  
|||  
Db 989 TATGAATATGAAA-----CTTCACTAGGAGCCCAAGATCCAA 1024  
317 SerProGlyProValGluIleLeuThr-----ThrValThrGluSerVal 331  
|||  
|||  
Db 1025 GAAGAGGAGAACAGGTGTGAATTAAGTGAAGGACGTCCAGTGAATATCACAGAACATATA 1084  
332 ThrGlyIleSerArgAsnValSerThrAsnValPhePheLysGlnHisAspTyrlIleIle 351  
|||  
|||  
Db 1085 ACCAAACTCTA-----TTTGTAAGTGGAGCTCACACTTTCAGACGGGA-----ATT 1132  
352 GluPhePheAspTyrlThrThrValLeuLysProSerLeuAsnPheThrAlaThrValLys 371  
|||  
|||  
Db 1133 CCCTTCTTTGGG-----CAG 1147  
372 ValThrArgAlaAspGlyAsnGlnLeuThrLeuGluLunArgArgAsnValValIle 391  
|||  
|||  
Db 1148 GTGGGCTCTAGTAGATGGGAAAGCGCTCCATATACA-----AATAAGTCATA 1195  
392 ThrValThrGlnArgAsnTyrlThrGluTyrlTrpSerGlySerAsnSerGlyAsnGlnLys 411  
|||  
|||  
Db 1196 TTCATC--AGAGAAATGAAGCAACTATTAC-----TCCAATGCTACCAGGAGTAC 1246  
412 MetGluAlaValGln--LysIleAsnTyrlThr--ValProGlnSerGlyThrPhe 428  
|||  
|||  
Db 1247 CATGGCCTTGTACAGTCTCTATCAACACACCAAGCTATGGTACCTCTTACTGTT 1306  
429 LysIleGluPhePro-----IleLeuGluAspSer 438  
|||  
|||  
Db 1307 AGGCGCATATTCAMAGATCGTAGTCCCTGTTACGGCTACCAAGTGGGCTGCAAGAAACAC 1366  
439 SerGluLeuGlnLeuLysAlaTyrlPheLeuGlySerLysSerSerMetAlaValHisSer 458  
|||  
|||  
Db 1367 GAAGAGCACATCCACTGCTTAT----- 1390  
459 LeuPheLysSerProSerLysThrTyrlGlnLeuLysThrArgAspGluAsnIleLys 478  
|||  
|||  
Db 1391 CTTGTGTCTTCCCAAGCAAGAGCTTTGTCCACTGAGCCCATGTCTCATGAACATACCC 1450  
479 ValGlySer-----ProPheGluLeuValValSerGlyAsnLys-----Arg 492  
|||  
|||  
Db 1451 TGTGGCCCTACTCAGACAGTCCAGGACCATTAATTCGATGATGGAGGACACCTGTGGGG 1510  
493 LeuLysGluLeuSer-----TyrMetValValSerArgGlyGlnLeuValAlaValGly 510  
|||  
|||  
Db 1511 CTGAAGAGCTCTCTCTTTTATTATCTATATATGGCAAAAGGAGGACATTGCCAAGCTGGG 1570  
511 -----LysGlnAsnSerThr-----MetPheSerLeuThr 520  
|||  
|||  
Db 1571 ACTCATGAGACTGCTGTGGAAGCAGGAACATGAAGGCCATTTTTCATCTCAATCCCT 1630  
521 ProGluAsnSerTrpThrProLysAlaCysValIleValTyrlGluAspAspGly 540  
|||  
|||  
::: |||  
::: |||

Db	1631	GTGAAGTCAGACATTGCTCCGTGGCTCGGTGGTTCGATCATTAAGCTGTTTACCTACCGG	1630
Qy	541	GlulleIleSerAspValLeuLysIleProValGlnLeuValPheLysAsnLysIleLys	560
Db	1691	GACCTGATTTGGGGATTCTGCAGAAAATATATGATGTGAAAATTGCTGGCCACAAAGGCTGAT	1750
Qy	561	LeuTyrTrpSerLysValLysAlaGlnProSerGluLysValSerLeuArgIleSerVal	580
Db	1751	TTGAGCTTCACGCCATCAACMAAGTCTCCAGGCTTCACACGCCCATCTCGAGTCAACCG	1810
Qy	581	ThrGlnProAspSerIleValGlyIleValAlaValAspLysSerValAsnLeuMetAsn	600
Db	1811	GCT---CCTCAGTCCGCTCGGGCCCTCCGCTGCTGTGAGCAACAAAGCTCTGTCTATAAG	1866
Qy	601	AlaSerAsnAspIleThrMetGlnAsnValValHisGluLeu--GluLeuTyrAsnThr	619
Db	1868	CCTGATGCTGAGCTCTCGGCGCTCTCGGTTTAACTGCTTACCTGACAGAAAGGACTCACT	1927
Qy	620	GlyTyr-----TyrLeuGlyMet	625
Db	1928	GGCTTCCCTGGGCTTTGAATGACAGAGCATGAAGCTGCATCAATCGTCAATAATGTC	1987
Qy	626	PheMetAsnSer-----PheAla	631
Db	1988	TATATTATATGGAATACATATATCTCAATATCAAGTACAAATGAAGAAGATATGTAACGC	2047
Qy	632	ValPheGlnGluCysGlyLeuTyrValLeuThrAspAlaAsnLeuThrLysAspTyrIle	651
Db	2048	TTCCATAGAGACATGCGTTTAAAGGATTCACCAACTCAAGATTCTGTAACCCAAATG	2107
Qy	652	AspGlyValIlyTrpAspAsnAlaGlu-----Tyr	660
Db	2108	TGTCACACGCTTCAACAGATGATGAATGATGATGAGACTGAAAGTCTACGTGATGTTTAT	2167
Qy	661	AlaGlnArgPheMetGluGlnAsnGluGlyHisIleValAspIleHisAspPheSerLeu	680
Db	2168	GAGTCACAGATGTAATGGGAAGAGCCATTCACGCCCTGGGACATGTGAAGAG-----	2218
Qy	681	GlySerSerProHis-----ValArgLysHisPheProGluThrTyrIleTyrLeu	697
Db	2219	-----CCTCACAGGAGACCGCTACGAAGAATCTTCCCTGACACATGATCTGGGAT	2266
Qy	698	AspThrAsnMetGlyTyrArgIleTyrGlnGluPheGluValThrValProAspSerIle	717
Db	2270	TTGGTGTGTGTAACTCAGACAGGGGTGGCTGAGTAGAGAGTAACAGTCCCTGACACACATC	2328
Qy	718	ThrSerTrpValAlaThrGlyPheValIleSerGluAspLeuGlyLeuGlyLeuThrThr	737
Db	2330	ACCGAGTGGGAAGGACGGGCGCTTCTCGCTGTCTGTAAGATGCTGACCTGTGATCTCTTCC	2388
Qy	738	ThrProValGlnLeuGlnAlaPheGlnProPhePheIlePheLeuAsnLeuProTyrSer	757
Db	2390	ACT---GGCTCTCTCCGAGCCCTTCACGCCCTTCTTTGGGAGCTTACAAATGCTTAACCT	2446
Qy	758	ValIleArgGlyGlnGluPheAlaLeuGluIleThrIlePheAsnTyrIleuLysAspAla	777
Db	2447	GTCGATCTGTGAGAGGCTTCACACTCAAGGCAAGGCTCTPAACTACTTCCCAATATGC	2506
Qy	778	ThrGluValLysValIleIleGlnLysSerAspLysPheAspIleLeuMetThrSerSer	797
Db	2507	ATCCGGGTCACTGTGCAGCTGGAACCCCTCTCCCGCTTCTTGTCTGTCCCAAGTAGAAG	2566
Qy	798	Glu-----IleAsnAlaThrGlyHisGlnGlnThrLeuLeuValProSer	812
Db	2567	GAAACAAGGCTCACTGATCTGTGTCAAAACGGGCGGCA-----	2606
Qy	813	GluAspGlyAlaThrValLeuPheProIleArgProThrHisLeuGlyGluIleProIle	832
Db	2606	-----ACGTGTCTCTGGGAGCAATACCCCAAGATCATTAGAAATGTGAATTTC	2655
Qy	833	ThrValAlaThrAlaLeuSer-----ProThrAla	841
Db	2654	ACTGTGACCGGACAGGACCTAGAGTCTCAAGAGCTGTGGACTAGAGTGGCTTTCATTT	2713

QY	842	-----SerAspAlaValThrGlnMetIleLeuValIysPalaGluGlyLe	856
DB	2714	CCTGAACACGGAAAGAAAGACACAGCTATCAAGCCTCTGTGGTTGAACCTGGAAAGACTA	2773
QY	857	GluIysSerIlyrSerGlnSerIleIleuLeuAspLeuThrAspAsnArgLeuGlnSerThr	876
DB	2774	GAGAAAGAAACACATTCACCTCCCTACTTTGTCCATCAGGTGGTGAGGTTCT------	2827
QY	877	LeuIlyrThrLeuSerPheSerPheProProAsnThrValThrGlySerGluArgValGln	896
DB	2828	--GAGAGATTATCCCTGAAACCTCCACCAAAATGTGTAGAAAGATCTGCCCGGACTTCT	2884
QY	897	IleThrAlaIleGlyAspValIleuGlyProSerIleAsnGlyLeuAlaSerLeuIleArg	916
DB	2885	GTCTCAGTTTGGGAGACATATTAAGGCTCTGCCATGCAAAACACACAAAAATCTTCTCCAG	2944
QY	917	MetProIlyrGlyCysGlyGluGlnAsnMetIleAsnPheAlaProAsnIleIyIleu	936
DB	2945	ATGCCCTATGGCTGTGGAGAGCAAAATATGCTCTTGTCTTAACTATATGTAAGTCTG	3004
QY	937	AspTyrIleuThrIlyLysLysGlnLeuThrAspAsnLeuLysGluIysPalaLeuSerPhe	956
DB	3005	GATTATCTAAATGAAACACAGCAGCTACTTCCAGAGGTCAAGTCCAAAGGCATTTGGCTAT	3064
QY	957	MetArgGlnGlyTyrGlnArgIleuLeuIlyrGlnArgIleuAspGlySerPheSerAla	976
DB	3065	CTCAACACTGGTTTCCAGACACAGTGTAACTTCAACATCAATGATGATGCTCTTCAAGCAC	3124
QY	977	PheGly---AsnTyrAspProSer---GlySerThrIleuSerAlaPheValLeuArg	994
DB	3125	TTTGGGAGCCATATGGCAGGACACAGGAGCAACCTGGCTCCACAGCCTTGTGTGTGAAG	3184
QY	995	CysPheLeuGlnAlaAspProIlyrIleAspIleAspGlnValLeuHisArgIlyrThr	1014
DB	3185	ACTTTTGGCCAGAGCTCGAGCTCATCTTATCATGATGAAGACACATTAACCAAGCCTC	3244
QY	1015	ThrTyrLeuIlysglyHisGlnLysSerAsnGlyGluPheTyrAspProGlyArgValIle	1034
DB	3245	ATATGGCTCTCCACAGGACAGAAAGAACATATGGCTGTTTACAGACCTCGGCTCATCTGCTC	3304
QY	1035	HisSerGluLeuGlnGlyIysAsnLysSerProValThrLeuThrAlaTyrIleValThr	1054
DB	3305	AACATATGCATTAAGGAGAGAGTGAAGATGAATGACCTCTCGGCTATATACCATC	3364
QY	1055	SerLeuLeuGlyTyrArgLysIyGlnProAsnIleAspValGlnGlnSerIleHisPhe	1074
DB	3365	GCCCTTCTGGAGATTCCTCTCAACAGTACATCACCTGTGTGCCAATGCCCTGTTTTGC	3424
QY	1075	LeuGlnSer-----GlnPheSerArgIlyIleSerAspAsnTyrThr	1088
DB	3425	CTGGAGTCAAGCTCGAAGACAGACAAAGAGGAGACCATGGC--AAGCAGATGATATATCC	3481
QY	1089	LeuAlaLeuIleThrTyrAlaLeuSerSerValGlySer---ProIlyPalaLysGluAla	1107
DB	3482	AAACACACTGCTGGCCTATGCTTTTGGCCCTGGCAAGTAAACAGACAAAGAGAAAGAAATA	3541
QY	1108	LeuAsnMetLeuThrTyrPargAlaGluGlnGluIlyGlyMetGlnPheTyr-----	1124
DB	3542	CTCAAGTCACTTAATGAGAAAGCTGTGAAGAAACAACTGTCTGCATTGGAGCGCCCT	3601
QY	1125	ValSerSerGlnSerIlyLysSerAspSerTyrGlnProArg-----SerLeuAspIle	1142
DB	3602	CAGAAACCCAGACCAAGCAGTGGGGCATTTTATACAAACCCAGAGCTCCCTGCTGAAGGTG	3661
QY	1143	GluValAlaAlaTyrAlaLeuLeuSerHisPheLeuGlnPheGln-----ThrSer	1159
DB	3662	GAGATGACATCTTATGTGCTCTGCT---TATCTCAAGGCCACAGCCACCAACCTCG	3718
QY	1160	GluGlyIlePro-----IleMetArgTyrLeuSerArgIlyArgAsnSerLeu	1175
DB	3719	GAGAGCCTGACCTTGCAACCAATGTGTAAAGTATCAAGAAAGCAGAAATGCCCAG	3778

```

Qy 1176 G1yGlyPheAlaSerThrGlnAspThrValAlaLeuValAlaLeuSerGlnPheAla 1195
Db 3779 GCGCGTCTCTCCACCCAGACAGACAGTGGTGGCTCCATGCTCTGTCCTCCAAATATGGA 3838
Qy 1196 AlaLeuMetAsnThrGlnArgThrAsnIleGlnValThrValThrGlnProSerSerPro 1215
Db 3839 GCCGCCACATTATACC--AGAGCTGGGAAGGCTGCACAGGAGCATTCACAGTCTTCAGGG 3895
Qy 1216 SerProLeuAlaValAlaGln----- 1222
Db 3896 ACATTTTCCAGCAATTCAGATGGACAAACAATGCGCTTACTGACAGAGTCTCA 3955
Qy 1223 -----PrometAlaValAsnIleSerAlaAsnGlyPheGlyPheAlaIleCys 1238
Db 3956 TTGCGCAGACTGCTGGGGAATACAGCATGAAGTGAAGTGAAGAGGATGTGTACTCTC 4015
Qy 1239 GlnLeuAsnValAlaThrAsnValAlaSerGlySerSerArgArgArgSerIle 1258
Db 4016 CAGACCTCTTGAATACATATAT-----CTC 4042
Qy 1259 GlnAsnGlnAlaAlaPheAspLeuAspValAlaValAlaLys-----GluAsnLysAsp 1275
Db 4043 CCAGAAAAGAAAGATTCCTTGGCTTTAGAGTGCAGACTGCGCTCAAACTTGAT 4102
Qy 1276 AspleuAsn--HisValAspLeuAsnValCysThrSerPheSerGlyProGly----- 1292
Db 4103 GAACCCAAAGCCCAACAGCTTCCAAATCTCCTAAGTGCAGTTACACAGGAGCCGC 4162
Qy 1293 ---ArgSerGlyMetAlaLeuMetGluValAsnLeuSerGlyPheMetValProSer 1311
Db 4163 TCTGCTCTCAACATGCGCATGTGATGTGAAGATGCTCTGCTTCATATCCCTTAAG 4222
Qy 1312 GluAlaIleSerLeu-----SerGlnThrValLysLysValGluTyAspHisGly 1328
Db 4223 CCAACAGTGAAGATTCCTTGAAGATCTAACATGTGAGCCGACAGAAATCAGACAGAAC 4282
Qy 1329 LysLeuAsnLeuTyLysAspSerValAsnGlnThrGlnPheCysValAsnIleProAla 1348
Db 4283 CATGCTTGAATTCCTTGAATGATGTGCAATCAGACACTGAGCTGATCTTCAAGGTT 4342
Qy 1349 ValArgAsnPheLysValSerAsnThrGlnAspAlaSerValSerIleValAspTyr 1368
Db 4343 CTGCAAGATGTCACAGTGAAGATCTCAACACCCATGAGAAAGCTATGATTTACTAC 4402
Qy 1369 GluProArgArgGlnAlaValAlaArgSerTyrAsnSer 1380
Db 4403 GAGACGATGAGTTTGCATCGCTGAGTACATGCT 4438

RESULT 13
US-10-076-816-56
; Sequence 56, Application US/10076816
; Publication No. US20030056244A1
; GENERAL INFORMATION:
; APPLICANT: Huang, Ning
; APPLICANT: Rodriguez, Raymond
; APPLICANT: Hagie, Frank E.
; TITLE OF INVENTION: Feed Additive Compositions and Methods
; FILE REFERENCE: 50665-8021.US00
; CURRENT APPLICATION NUMBER: US/10/076,816
; PRIOR APPLICATION NUMBER: US 60/269,188
; PRIOR FILING DATE: 2001-02-14
; PRIOR APPLICATION NUMBER: US 09/847,232
; PRIOR FILING DATE: 2001-05-02
; PRIOR APPLICATION NUMBER: US 60/266,929
; PRIOR FILING DATE: 2001-02-06
; PRIOR APPLICATION NUMBER: US 60/201,182
; PRIOR FILING DATE: 2000-05-02
; NUMBER OF SEQ ID NOS: 60
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 56
; LENGTH: 4577
; TYPE: DNA

```

```

; ORGANISM: Homo sapiens
US-10-076-816-56
Alignment Scores:
Pred. No.: 3,566-133
Score: 1440.00
Percent Similarity: 46.20%
Best Local Similarity: 29.06%
Query Match: 15
DB: 15
Gaps: 59
Length: 4577
Matches: 451
Conservative: 266
Mismatches: 551
Indels: 284
Gaps: 59

US-10-020-095-4 (1-1428) x US-10-076-816-56 (1-4577)
Qy 12 LeuLeuCysValCysThrAlaAlaLeuAlaValAlaProGlyProArgPheLeuValThr 31
Db 86 CTCCTGGTCTCTCCGCCACAGAGCGCTCAGTCTCTGAAAACCGCAGATATAGTTCTTG 145
Qy 32 AlaProGlyLysIleLeuArgProGlyGlyAsnValThrIleGlyValAlaLeuLeuGluHis 51
Db 146 GTCCCTCTCTGCTCCACT--GAGACCACTGAGAGAGGCTGTGCTCTTGAAGCTAC 202
Qy 52 CysProSerGlnValThrValLysAlaGluLeuLeuLysThrAlaSerAsnLeuThrVal 71
Db 203 CTGAATGACAGACTGACTGTAACTGCTTCTTGAAGTCTGTCAAGGAAACAGAGACCTC 262
Qy 72 ---SerValLeuGluAlaGluGly---ValPheGluLysGlySerPheLysThrLeuThr 89
Db 263 TTCACGTACCTGAGGCGGAGATGACGTATCTCACTGTGTGCGCTTC----- 310
Qy 90 LeuProSerLeuProLeuAsnSerAlaAspGlu-----IleTyGluLeuAlaGlyAl 106
Db 311 -----GCTGTCCCAAGCTTTCATCCAAATGAGAGATATGTTCTTCACTGCAAGG 364
Qy 107 ThrGlyArgThrGlnAspGluIleLeuPheSerAsnSerThrArgLeuSerPheGluThr 126
Db 365 AAAGAACCAACCCAAAGAA-----TTTAAAGCCGAGCACAGTGAATGCTTAAGAAC 415
Qy 127 LysArgIleSerValPheIleGlnThrAspLysAlaLeuTyLysProLysGlnGluVal 146
Db 416 GAGGACAGTGTGCTCTTGTCCAGACAGCAAAATCAATCAACAAACAGGAGCAGACGTG 475
Qy 147 LysPheArgIleValThrLeuPheSerAspPheLysProTyrLysThrSerLeuAsnIle 166
Db 476 AAATTCGTGTGCTTCCATGATGAAATCTTCAACCCCTGAAATGATGATTTCACATA 535
Qy 167 Leu---IleLysAspProLysSerAsnLeuIleGlnGlnThrLeuSerGlnGlnSerAsp 185
Db 536 GTATACATTCAGATCCCAAGAGAAATCCGATGCCAATGGCAGAGTTTCCAGTTAGAG 595
Qy 186 LeuGlyValIleSerLysThrPheGlnLeuSerSerHisProIleLeuGlyAspTyrSer 205
Db 596 GGTGGCTCAAGAAATTTCTTTCCCTCTATCAGAGCCCTTCCAGGGCTCTCAAG 655
Qy 206 IleGlnValGlnValAsnAsp-----GlnThrTyThrGlnSerPheGlnValSerGlu 223
Db 656 GTGTGTGTACAGAAATCAGGTGAGAGACAGACAC---CCTTTCACCGTGAGGAA 712
Qy 224 TyValLeuProLysPheGluValThrLeuGlnThrProLeuTyrcysSerMetLeuAsn 243
Db 713 TTGTGCTTCCCAAGTTGAAGTAAAGTAAACAGTGAAGTGAAGTAAATCAACATCTTGA 772
Qy 244 LysIleLeuAsnGlyThrIleThrAlaLysTyThrTyGlyLysProValLysGlyAsp 263
Db 773 GAGAGATGAATATACAGTGTGTGCGCTTATACATATGGAAGGCTGTCCCTGAGCAT 832
Qy 264 ValThrLeuThrPheLeuProLeu----- 271
Db 833 GTGACTGTAGCATTTGCAGAAAGTATAGTACGCTTCCGACTGCCAGCGTGAAGATTC 892
Qy 272 -----SerPheTyrGlyLys 276
Db 893 CAGGCTTCTGTGAGAAATTCAGTGAACAGCTTAAACAGCCAGCGCTTCTTATACAGCA 952

```



QY 277 LysLeuSerHisIleThrIleThrPheLysIleAsnGlySerAlaAsnPheSerPheAsnAsp 296  
 DB 953 GTAAAA-----ACCAAGGCTTCCAGCAAGAGG-----AAGAG 988  
 QY 297 GluGluMetLeuSerAsnValMetLeuSerSerSerAsnGlyLeuSerGluIleuLeuSer 316  
 DB 989 TATGAATATGAA-----CTTCACACTGAGGCCAGATCCAA 1024  
 QY 317 SerProGlyProValGluIleLeuThr-----ThrValThrGluSerVal 331  
 DB 1025 GAAGAAGAAACAGATGCTGGAAATGACTGGAAAGCAGTCCAGTAATTCACAGAACCATTA 1084  
 QY 332 ThrGlyIleSerArgAsnValSerThrAsnValPhePheLysGlnHisAspTyrIleIle 351  
 DB 1085 ACCAAACTCTCA---TTGTGAAGTGAATGACACTTTCGACAGGGA-----ATT 1132  
 QY 352 GluPhePheAspTyrThrThrValLeuLysProSerLeuAsnPheThrAlaThrValLys 371  
 DB 1133 CCCTTCTTGGG-----CAG 1147  
 QY 372 ValThrArgAlaAspGlyAsnGlnLeuThrLeuGluGluArgAsnAsnValIle 391  
 DB 1148 GTCCGCTAGTAAATGGAAGAGCGTCCCTATACCA-----ATAAAGTCATA 1195  
 QY 392 ThrValThrGluArgAsnTyrThrGluTyrTyrSerGlySerAsnSerGlyAsnGlnLys 411  
 DB 1196 TTCATC---AGAGAAATGAAGCAAACTATTAC---TCCAATGCTACCAAGCATGAG 1246  
 QY 412 MetGluAlaValGln---LysIleAsnTyrThr---ValProGlnSerGlyThrPhe 428  
 DB 1247 CATGCCCTGTATACATCTCTATACACCAACCATGATGGTACCTCTCTTACTGTT 1306  
 QY 429 LysIleGluPhePro-----LleuGluAspSer 438  
 DB 1307 AGGGCTAATTACAGATCTCTACTCCCTGTACGCGTCAACAGTGGGTGTCAGAGAACAC 1366  
 QY 439 SerGluLeuGlnLeuLysAlaTyrPheLeuGlySerLysSerSerMetAlaValHisSer 458  
 DB 1367 GAAAGGCAATCACTGCTTAT----- 1390  
 QY 459 LeuPheLysSerProSerLysThrTyrIleGlnLeuLysThrArgAsnGluAsnIleLys 478  
 DB 1391 CTGTGTCTTCCCCAAGAGAGCTTGTCCACCTTGAGCCCATGCTCATGAATACCC 1450  
 QY 479 ValGlySer-----ProPheGluLeuValValSerGlyAsnLys-----Arg 492  
 DB 1451 TGTGGCCATATCAGACAGTCCAGGACATTAATATCTGATAGAGGACACCTGCTGGG 1510  
 QY 493 LeuLysGluLeuSer-----TyrMetValValSerArgGlyGlnLeuValAlaValGly 510  
 DB 1511 CTGAAGAAGCTCTCTTTTATTTATCTGATATATGCAAGAGGAGGACTGTGCCAAGCTGG 1570  
 QY 511 -----LysGlnAsnSerThr-----MetPheSerLeuThr 520  
 DB 1571 ACTCATGAGCTGCTGTGGAAGCAGAAACATGAGAGGCGCATTTTTCATCTCATATCCCT 1630  
 QY 521 ProGluAsnSerThrProLysAlaCysValIleValTyrTyrIleGluAspAspGly 540  
 DB 1631 GTGAAGTCAGACATTCCTCTGTGCTGGTGTGCTCATATGCTGTATTAATCTTACCTGCG 1690  
 QY 541 GluIleIleSerAspValLeuLysIleProValGlnLeuValPheLysAsnLysIleLys 560  
 DB 1691 GAGCTGATTTGGGATTCGTGCAAAATATGATGTTGAAAATGTGTGGCCAAAGAGTGAT 1750  
 QY 561 LeuTyrTyrSerLysValLysAlaGluProSerGluLysValSerLeuArgIleSerVal 580  
 DB 1751 TTGAGCTTCAGCCCATCAAAAGTCTCCAGCTTCACAGCCCATCGCAGATCACAGCG 1810  
 QY 581 ThrGlnProAspSerLysIleValGlyIleValAlaValAspLysSerValAsnLeuMetAsn 600  
 DB 1811 GCT---CCTAAGTTCCTGCGCCCTGCTGCTGAGCCAAAGCGTGTGCATGAAG 1867  
 QY 601 AlaSerAsnAspIleThrMetGluAsnValValIleGluLeu---GluLeuTyrAsnThr 619

DB 1868 CCTGATGCTAGCTCTCGCGCTCTGCTTATCAACCTGTATACAGAAAGACTCTCACT 1927  
 QY 620 GlyTyr-----TyrLeuLysMet 625  
 DB 1928 GGGCTTCCTGGGCTTTGAATGACCAAGACATGAAGACTGATCATATCTATATATGCT 1987  
 QY 626 PheMetAsnSer-----PheAla 631  
 DB 1988 TATATTAATGAAATCATATATCTTCAGTATCAAGTACAAATGAAAAGATATGTACAC 2047  
 QY 632 ValPheGlnGluCysGlyLeuTyrValLeuThrAspAlaAsnLeuThrLysAspTyrIle 651  
 DB 2048 TTCCTAGAGACATGGGCTTAAAGCAATTCACCACTCAAGATTCGTAAACCAAAATG 2107  
 QY 652 AspGlyValTyrAspAsnAlaGlu-----Tyr 660  
 DB 2108 TGTCCACAGCTTCAACAGTATGAATGACATGAGCTGAAGCTTACGTATAGTTTAT 2167  
 QY 661 AlaGluArgPheMetGluGluAsnGluGlyHisIleValAspIleHisAspPheSerLeu 680  
 DB 2168 GAGTACAGATGTATGGAAGAGGCAATGACAGCGCTGTGATTTGAAGAC----- 2218  
 QY 681 GlySerSerProHis-----ValArgLysHisPheProGluThrTyrIleThrPhe 697  
 DB 2219 -----CTCACAGGAGACCGTACGAAAGTACTTCCCTGAGACATGATCTGGAT 2269  
 QY 698 AspThrAsnMetGlyTyrArgIleTyrGlnGluPheGluValThrValProAspSerIle 717  
 DB 2270 TTGGTGTGTGTAATCACTACAGAGGGGTGCTGATGAGATGAACGTCCATGACCATC 2329  
 QY 718 ThrSerThrValAlaThrGlyPheValIleSerGluAspLeuGlyLeuGluThrThr 737  
 DB 2330 ACCGATGGAAGGCGAGGCTTCTGCTGTGTGAAGATGCTGGACTTGATATCTTCC 2389  
 QY 738 ThrProValGluLeuGlnAlaPheGlnProPhePheIlePheLeuAsnLeuProTyrSer 757  
 DB 2390 ACT---GCCTCTCTCGAGCCTTCCAGCCCTTCTTGTGGAGCTTACAAATGCTTACTCT 2446  
 QY 758 ValIleArgGlyGluGluPheAlaLeuGluIleThrIlePheAsnTyrLeuLysAspAla 777  
 DB 2447 GTGATTTCTGGAGAGGCTTTCACACTCAAGGCGCAAGCTTAACTTCCCAATATGC 2506  
 QY 778 ThrGluValLysValIleIleGluLysSerAspLysPheAspIleLeuMetThrSerSer 797  
 DB 2507 ATCCGAGTCAATGTGCTAGCTGAACTCTCCGCTTCTTGTGCTGCCAGTGAAG 2566  
 QY 798 Glu-----IleAsnAlaThrGlyHisGlnGlnThrLeuValProSer 812  
 DB 2567 GAAACAAGCGCTCAGCTGCATCTGTGCAAAAGGCGGCA----- 2605  
 QY 813 GluAspGlyAlaThrValLeuPheProIleArgProThrHisLeuGlyGluIleProIle 832  
 DB 2606 -----ACTGTCTCTGGGCAGTAAACCCCAAGTCAATTAGAAATGTGAATTTC 2653  
 QY 833 ThrValThrAlaLeuSer-----ProThrAla 841  
 DB 2654 ACTGTAGGCGCAGAGCACTAGAGTCTCAAGAGCTGTGGAGCTGAGGTGCTTCAGTT 2713  
 QY 842 -----SerAspAlaValThrGlnMetIleLeuValLysAlaGluLysIle 856  
 DB 2714 CCTGAACAAGAAAGAAAGACACACTCATCAAGCTCTGTGTGTAACCTGAAGACATTA 2773  
 QY 857 GluLysSerTyrSerGlnSerIleLeuLeuAspLeuThrAspAsnArgLeuGlnSerThr 876  
 DB 2774 GAAAGAAACAAACATTAACCTCTTACTTGTCCATATGATGATGAGATTTCT----- 2827  
 QY 877 LeuLysThrLeuSerPheSerPheProProAsnThrValThrGlySerGluArgValGln 896  
 DB 2828 ---GAAAGATTAATCTTGAACCTGCAACCAATGTGTGAAGAAATGTGCCAGACTTCT 2884  
 QY 897 IleThrAlaIleGlyAspValLeuGlyProSerIleAsnGlyLeuAlaSerLeuIleArg 916

```

Db      2885 GTCTCAGTTTGGGAGACATATTAGGCTCTGCCATGCAAAAACAACAATCTTCTCCAG 2944
Qy      MetProtyrGlyCySGlyGluGlnAsnMetIleasnPheAlaProAsnIleTyrIleu 936
Db      2945 ATGCCCTATGGCTGTGGAGACAGAAATATGCTCTTGTCTCTTAACATATATGTA 3004
Qy      AspTyrLeuThrIleValGlyGlnLeuThrAspAsnLeuValGlyValIleuSerPhe 956
Db      3005 GATTATCTTAATTAAGAAACACAGACCTTACTCCAGAGCTCAAGTCCCAAGGCATTTGGCAT 3064
Qy      MetArgGlnGlyTyrGlnArgGlnLeuLeuTyrGlnArgGlnAspGlySerPheSerAla 976
Db      3065 CTCAACACTGGTATGAGACAGAGCTTGAATTAACAACATATGATGCTCTTCAACAGACC 3124
Qy      PheGly---AsnTyrAspProSer---GlySerThrTrpLeuSerAlaPheValLeuArg 994
Db      3125 TTTGGGAGGCGATATGGCAGGAAACCGGCAACCTGTGCTCACAGCCTTTGTTCTAGAG 3184
Qy      CysPheLeuGlnAlaAspProTyrIleAspIleAspGlnAsnValLeuIleAspThrTyr 1014
Db      3185 ACTTTGGCCCAAGCTCGAGCCTTACTTCATCGATGAAGCAACATTAACCCAGCCCTC 3244
Qy      ThrTrpLeuValGlyHisGlnLeuSerAsnGlyGlnPheTrpAspProGlyArgValIle 1034
Db      3245 ATATGGCTCTCCAGAGGCAAGACCAATAGCTGTTCAGAGACTCTGGGTCACTGCTC 3304
Qy      HisSerGlnLeuGlnGlyGlyAsnLeuSerProValThrLeuThrAlaTyrIleValThr 1054
Db      3305 AACATGTCATTAAGGAGAGAGCTAGAAAGATGAAGACCTCTCGGCTATATACCAATC 3364
Qy      SerLeuLeuGlyTyrArgIleTyrGlnProAsnIleAspValGlnGluSerIleHisPhe 1074
Db      3365 GCCCTTGGAGATTCCTCTCAAGTCACTCACTGCTGTGTCGCAATGCCCTGTTTTCG 3424
Qy      LeuGluSer-----GluPheSerArgGlyIleSerAspAsnTyrThr 1088
Db      3425 CTGAGTGAGCCTGAGACAGACACAAAGAGGAGCCATGGC---AGCCATGTATATAC 3481
Qy      LeuAlaLeuIleThrTyrAlaLeuSerSerValGlySer---ProValAlaValGlyAla 1107
Db      3482 AAAGCATGCTGGCTATATGCTTTCCTGCGCAGGTAAACAGACAGACAGAGAGAAAGTA 3541
Qy      LeuAsnMetLeuThrTrpArgAlaGlnGlnGlyMetGlnPheTrp----- 1124
Db      3542 CTCAAGTCACTTAATGAGAGAGCTGTGAAGAAACACTGTGTCATTTGGAGCGCCCT 3601
Qy      ValSerSerGluSerIleValSerAspSerTrpGlnProArg-----SerLeuAspIle 1142
Db      3602 CAGAAACCCAGAGGACCCAGTGGGCGATTTTTCAGAACCCCGAGCTCCTCTGCTGAGG 3661
Qy      GluValAlaIleTyrAlaLeuLeuSerHisPheLeuGlnPheGln-----ThrSer 1159
Db      3662 GAGATGACATCTTATGCTCTCTGCTGCT---TATCTCAGCGGCCCAAGCCCAACTTCG 3718
Qy      GluGlyIlePro-----IleMetArgTrpLeuSerArgIleArgAsnSerLeu 1175
Db      3719 GAGGACCTGACCTGCAACCAACATGCTGAAGTGAAGTGAAGAGAGCAACAATATGCCAG 3778
Qy      GlyGlyPheAlaSerThrGlnAspThrThrValAlaLeuValAlaLeuSerGlnPheAla 1195
Db      3779 GGCCTGTTCTCTCCACCCAGAGACACAGTGTGCTCCACATGCTGTCTCCAAATATATGCA 3838
Qy      AlaLeuMetAsnThrGlnArgThrAsnIleGlnValThrValThrGlyProSerSerPro 1215
Db      3839 GCCCGCACATTATAC---AGGACTGGGAAGGCTGCAAGGAGTACTTCCAGTCTTCAAGG 3895
Qy      SerProLeuAlaValAlaGln----- 1222
Db      3896 ACATTTTCCAGCAATTCAGAGTGAACAACAATGCGCTTACTGACGACAGTCTCA 3955
Qy      1223 -----PrometAlaValAsnIleSerAlaAsnGlyPheGlyPheAlaIleCys 1238
Db      3956 TTGCGACAGCTGCTGGGGAATACAGCATGAAGTGAAGAGATGTGTCTACTTC 4015

```

```

Qy      1239 GlnLeuAsnValValTyrAsnValIleValSerGlySerSerArgArgArgSerIle 1258
Db      4016 CAGACCTCTCTGAAATATCAATATT-----CTC 4042
Qy      1259 GlnAsnGlnGlnAlaPheAspLeuAspValAlaValIle-----GluAsnIleAsp 1275
Db      4043 CAGAAAGAGAAAGTCCCTTGTGCTTGAAGAGTCAAGCTGCTGCTCAAACTTGAT 4102
Qy      1276 AspLeuAsn---HisValAspLeuAsnValCysThrSerPheSerGlyProGly----- 1292
Db      4103 GAACCCAAAGCCACACACAGCTTCCAAATCTCCCTTAAGTCTAGTACAGAGGAGCCG 4162
Qy      1293 ---ArgSerGlyMetAlaLeuMetGluValAsnLeuLeuSerGlyPheMetValProSer 1311
Db      4163 TCTGCTCCCAATGCGGATCGTGAATGTGAAGATGCTGCTGCTTCAATTCCTGCAAG 4222
Qy      1312 GluAlaIleSerLeu-----SerGluThrValIleValValGlyValGlyTyrAspHisGly 1328
Db      4223 CCAACAGTGAATATGCTTGAAGATCTAACCATGATGAGCCGACAGAGATCAGACCAAC 4282
Qy      1329 LysLeuAsnLeuTyrIleAspSerValAsnGluThrGlnPheCysValAsnIleProAla 1348
Db      4283 CATGCTTGAATTAATCTTGAATAGTGTCAAAATCAACACACTGAGCTTGTCTTCAAGGT 4342
Qy      1349 ValArgAsnPheLysValSerAsnThrGlnAspAlaSerValSerIleValAspTyrTyr 1368
Db      4343 CTCAAGATGTCCACAGTAAAGATCTCAAAACCGCATGTGAAGTGAATGATTAATCTAC 4402
Qy      1369 GluProArgArgGlnAlaValArgSerTyrAsnSer 1380
Db      4403 GAGACGATGATGTTTCAATCGCTGAGTACAAATGCT 4438

```

## RESULT 14

```

US-10-240-965-178
: Sequence 178, Application US/10240965
: Publication No. US20030165924A1
: GENERAL INFORMATION:
: APPLICANT: INCYTE GENOMICS, INC.
: APPLICANT: SHIFFMAN, DOV
: APPLICANT: SOMOGYI, ROLAND
: APPLICANT: LAMN, RICHARD M.
: APPLICANT: SEILHAMER, JEFFREY J.
: APPLICANT: PORTER, GORDON J.
: APPLICANT: MIKITA, THOMAS
: APPLICANT: TAL, JULIE
: TITLE OF INVENTION: GENES EXPRESSED IN FOAM CELL DIFFERENTIATION
: FILE REFERENCE: PA-0025 PCT
: CURRENT APPLICATION NUMBER: US/10/240,965
: CURRENT FILING DATE: 2002-10-04
: PRIOR APPLICATION NUMBER: 60/195,106
: PRIOR FILING DATE: 2000-04-05
: NUMBER OF SEQ ID NOS: 276
: SOFTWARE: PERL Program
: SEQ ID NO 178
: LENGTH: 4809
: TYPE: DNA
: ORGANISM: Homo sapiens
: FEATURES:
: NAME/KEY: misc feature
: OTHER INFORMATION: Incyte ID No. US20030165924A1 1000033.6
US-10-240-965-178

```

## Alignment Scores:

```

Pred. No.: 3,87e-133 Length: 4809
Score: 1440.00 Matches: 451
Percent Similarity: 46.20% Conservative: 266
Best Local Similarity: 29.06% Mismatches: 551
Query Match: 19,60% Indels: 284
DB: Gaps: 59

```

US-10-020-095-4 (1-1428) x US-10-240-965-178 (1-4809)

QY 12 LeuLeuCyValCyThrAlaAlaValAlaProGlyProArgPheLeuValThr 31  
DB 116 CTCTGTGCTCTCCGCCACAGACGCCCTCAGTCTCGAATAACCGCATATATGTTCTG 175  
QY 32 AlaProGlyIleIleArgProGlyAlaValThrIleGlyValGluLeuGluHis 51  
DB 176 GTCCCTCTCCCTGCTCACACT--GAGACCACTGAGAGAGGCTGTCTCTTCGAGCTAC 232  
QY 52 CysProSerGlnValThrValIleValAlaGluLeuLeuLeuThrAlaSerAsnLeuThrVal 71  
DB 233 CTGAATGAGACAGTGACTGTAAGTGTCTTCTGAGTCTGTACGGGAAACAGAGCCCTC 292  
QY 72 ---SerValLeuGluAlaGluGly--ValPheGluGlySerPheLeuThr 89  
DB 293 TTTCAGTACCTGTGAGCGGAGATGACGTATCCACTGTGCGCTTC----- 340  
QY 90 LeuProSerLeuProLeuAsnSerAlaAspGlu-----IleGlyGluLeuArgVal 106  
DB 341 -----GCTGCCCAAGTCTTCATCCATGAGAGATATGTTCTCCTCACTGTCCAAAGTG 394  
QY 107 ThrGlyArgThrGlnAspGluIleLeuPheSerAsnSerThrArgLeuSerPheGluThr 126  
DB 395 AAGAGCAACCAAGCA-----TTTAAAGCGGACCAACAGTGAAGTTAAGAAC 445  
QY 127 LysArgIleSerValPheIleGlnThrAspIleValAlaLeuTyrlsPheProGlyGlnGluVal 146  
DB 446 GAGCAGACTGTGCTTGTTCAGACAGCAATCAATCAATCAACAGGACAGAGT 505  
QY 147 LysPheArgIleValThrLeuPheSerAspPheLeuProTyrlsThrSerLeuAsnIle 166  
DB 506 AATATTCGTTGTTCTCCATGATGATAAATTTCACCCCTGATGATGATGATTCACATA 565  
QY 167 Leu---IleLysAspProLysSerAsnLeuIleGlnGlnThrLeuSerGlnIleAsp 185  
DB 566 GATATCATTACAGATTCACAAAGAAATCCCATGACATGACAGAGATTTCCAGTTAGAG 625  
QY 186 LeuGlyValIleSerLysThrPheGlnLeuSerSerHisIleProIleLeuGlyAspTrpSer 205  
DB 626 GGTGGCTCAAGCAATTTCTTTTCCCTCATGAGAGCCCTTCAGAGGCTCTACAG 685  
QY 206 IleGlnValGlnValAsnAsp-----GlnThrTyrlsGlnSerPheGlnValSerGlu 223  
DB 686 GTGCTGTATCAGAAAGAAATCAGGTGAGAGACAGAGAC-----CTTTACCCGTGAGAGAA 742  
QY 224 TyrValLeuProLysPheGluValThrIleGlnThrProLeuTyrlsSerMetLeuAsn 243  
DB 743 TTGTGTTCTCCAAAGTTGAAAGTACAAAGTACAGTCCAAAGATTAATCACTTTGAA 802  
QY 244 LysHisIleLeuAsnGlyThrIleThrAlaLysTyrlsThrTyrlsGlyProValIleGlyAsp 263  
DB 803 GAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 862  
QY 264 ValThrLeuThrPheLeuProLeu----- 271  
DB 863 GTGACTGTGACATTGTCAGAAAGTATGATGACGTTCCGACTGACACGCTGAAGATTCA 922  
QY 272 -----SerPheTrpGlyLys 276  
DB 923 CAGGCTTTCTGTGAGAAATTCAGTGAAGCTTAACAGCAGCATGCTGCTTATACAGCA 982  
QY 277 LysLysAsnIleThrLysThrPheLysIleAsnGlySerAlaAsnPheSerPheAsnAsp 296  
DB 983 GTAATA-----ACCAAGCTTCTCCAGCTGAAGAG-----AAGAG 1018  
QY 297 GluGluMetLysAsnValMetAspSerSerAsnGlyLeuSerGluTyrlsLeuAspLeuSer 316  
DB 1019 TATGAATGAAA-----CTTCACTGAGGCGCCAGATCCAA 1054  
QY 317 SerProGlyProValGluIleLeuThr-----ThrValThrGluSerVal 331  
DB 1055 GAAAGAGAAACAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1114  
QY 332 ThrGlyIleSerArgAsnValSerThrAsnValPhePheGlnHisAspTyrlleIle 351

DB 1115 ACCAAACCTCA-----TTTGTGAAGTGACCTCACACTTTCAGACAGGGA-----ATT 1162  
QY 352 GluPhePheAspTyrlThrThrValLeuLysProSerLeuAsnPheThrAlaThrValLys 371  
DB 1163 CCTTCTTTGGG-----CAG 1177  
QY 372 ValThrArgAlaAspGlyAsnGlnLeuThrLeuGluGluArgArgAsnValValIle 391  
DB 1178 GTGCCCTAGTATGATGAGAAAGCGCTCCCTATACA-----AATTAAGTCA 1225  
QY 392 ThrValThrGlnAsnThrTyrlGluTyrlsTrpSerGlySerAsnSerGlyAsnGlnLys 411  
DB 1226 TTATC--AGAGAAATGAAGCAAACTATAC-----TCCAATGCTACACGAGATGAG 1276  
QY 412 MetGluAlaValGln--LysIleAsnTyrlThr-----ValProGlnSerGlyThrPhe 428  
DB 1277 CATGCCCTTGATACCTTCTATCAACACCAATGTTATGAGTACCTCTTATCTGTT 1336  
QY 429 LysIleGluPhePro-----IleLeuGluAspSer 438  
DB 1337 AGGTCATTATACAGATCGTATGCCCTTACAGGCTACAGTGGGTCCAGAAAGAAC 1396  
QY 439 SerGluLeuGlnLeuLysAlaTyrlPheLeuGlySerLysSerMetAlaValHisSer 458  
DB 1397 GAAGAGCAGATCACTGCTTAT----- 1420  
QY 459 LeuPheLysSerProSerLysThrTyrlIleGlnLeuLysThrArgAspGluAsnIleLys 478  
DB 1421 CTGTGTTCTCCCAAGCAAGCTTGTCCACCTTGAGCCCATGTCTCATGAACTACCC 1480  
QY 479 ValGlySer-----ProPheGluLeuValValSerGlyAsnLys-----Arg 492  
DB 1491 TGTGCCATCTACAGACGTCCAGGACATATATTTGATGAGAGGACCCCTGCTGGGG 1540  
QY 493 LeuLysGluLeuSer-----TyrMetValValSerArgGlyGlnLeuValAlaValGly 510  
DB 1541 CTGAAGAGACTCTCTCTTATATATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1600  
QY 511 -----LysGlnAsnSerThr-----MetPheSerLeuThr 520  
DB 1601 ACTCATGACGTCTGTGTAACAGAGACATGAAAGGCCATTTTCCATCTCAATCCCT 1660  
QY 521 ProGluAsnSerTrpThrProLysAlaCysValIleValTyrlTyrlIleGluAspAspGly 540  
DB 1661 GTGAAGTGAACATTTGCTCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1720  
QY 541 GluIleIleSerAspValLeuLysIleProValGlnLeuValPheLysAsnLysIleLys 560  
DB 1721 GACGTGATTTGGGATTTGCAAAATATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1780  
QY 561 LeuTyrlsPheLysValLysAlaGluProSerGluLysValSerLeuArgIleSerVal 580  
DB 1781 TTGAGCTTCAAGCCCATCAAAAGCTCCCAAGCTCCACAGCCCACTGCGAGTCAAGCG 1840  
QY 581 ThrGlnProAspSerIleValGlyIleValAlaValAspLysSerValAsnLeuMetAsn 600  
DB 1841 GCT--CCTCAGTCCGTGTGCGCCCTCGTCTGTGAGCAAAAGCGTCTGCTCATGAG 1897  
QY 601 AlaSerAsnAspIleThrMetGluAsnValValIleGluLeu-----GluLeuTyrlAsnThr 619  
DB 1898 CCGATGCTGAGCTCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1957  
QY 620 GlyTyrl-----TyrLeuGlyMet 625  
DB 1958 GCGTCCCTGCGGCTTGAATGACAGGACGATGAGACTGATCAATCTCATATATCA 2017  
QY 626 PheMetAsnSer-----PheAla 631  
DB 2018 TATATTAATGAAATCAATATATCTCAAGTATCAAAATGAAAGATATGATGATGATGATGATGAT 2077  
QY 632 ValPheGlnGluCysGlyLeuTrpValLeuThrAspAlaAsnLeuThrLysAspTyrlle 651

Db 2078 TTCTAGAGACATGGGCTTAAGGCAATTCAACAATCTCAAGATTCTGTAACCCCAAAATG 2137  
 Qy AspGlyValTyrAspAsnAlaGlu-----Tyr 660  
 Db 2138 TGTCCACAGCTTCAACAGTATGAATGATGAGCCTGAGCTACGTCGTTTAT 2197  
 Qy AlaGluArgPheMetGluGluAsnGluGlyHisLeuAlaAspIleHisAspPheSerLeu 680  
 Db 2198 GAGTCAGATGTAAATGGGAAGAGCCATGACGCTGTCGATGTTGAAG- 2248  
 Qy GlycerSerProHis-----ValArgLysHisAspPheProGluThrTrpIleTrpLeu 697  
 Db 2249 -----CCTCACACGAGACCGTACGAAAGTACTTCCCTGAGACATGATCTGGGAT 2299  
 Qy AspThrAsnMetGlyTyrArgIleTyrGlnGluPheGluValThrValProAspSerIle 717  
 Db 2300 TTGGTGGTGTAAATCTCAGCAGGTGTGCTGATGAGTAAAGTAAAGTCCCTGACACCATC 2359  
 Qy 718 ThrSerTrpValAlaThrGlyPheValIleSerGluAspLeuGlyLeuThrThr 737  
 Db 2360 ACCGAGTGAAGAGGAGGCGCTTCTGCTGTGTAAGATGCTGGACTGTATCTCTCC 2419  
 Qy 738 ThrProValGluLeuGlnAlaPheGlnProPheHeliPheLeuAsnLeuProTyrSer 757  
 Db 2420 ACT---GCCCTCTCCGAGCCTTCCAGCCTTCTTGTGAGCTCACAAATCCCTTACTCT 2476  
 Qy 758 ValIleArgGlyGluGluPheAlaLeuGluIleThrIlePheAsnTyrLeuLysAspAla 777  
 Db 2477 GTGATTCGTGAGAGGCTTACACTCAAGGCCACAGCTCTTAACTCTTCCCAATGTC 2536  
 Qy 778 ThrGluValLysValIleIleGluLysSerAspLysPheAspIleLeuMetThrSerSer 797  
 Db 2537 ATCCGGGTCACTGTGACAGCTGGAAGCCTCCCGCCTTCTTACGTCGCCAGTGAAG 2596  
 Qy 798 Glu-----IleAsnAlaThrGlyHisGlnGlnThrLeuValProSer 812  
 Db 2597 GAACAAGCGCTCACTGCATCTGTGCAACCGGCGCA----- 2635  
 Qy 813 GluAspGlyAlaThrValLeuPheProIleArgProThrHisLeuGlyGluIleProIle 832  
 Db 2636 -----ACTGTCTCTGGGACGTAAACCCCAAGTCAATTAGAAATGTAAATTTC 2683  
 Qy 833 ThrValThrAlaLeuSer-----ProThrAla 841  
 Db 2684 ACTGTAGCGCAGAGGACACTGAGTCTCAAGAGCTGTGGAGCTGAGTGCCTTCAATT 2743  
 Qy 842 -----SerAspAlaValThrGlnMetIleLeuValLysAlaGluGlyIle 856  
 Db 2744 CCTGAACAAGAGAGAAAGACACAGTCAAGCTCTGTGTGTAACCTGAAGAGACTA 2803  
 Qy 857 GluLysSerTyrSerGlnSerIleLeuLeuAspLeuThrAspAsnArgLeuGlnSerThr 876  
 Db 2804 GAGAAAGAAACAACTTCACTCCCTACTTGTTCATCAGGTGTGAGTTTC----- 2857  
 Qy 877 LeuLysThrLeuSerPheSerPheProProAsnThrValThrGlySerGluArgValGln 896  
 Db 2858 ---GAGAATATTATCCGTAACCTGCCACAAATGTGTGAGAAGATCTGCCGAGCTTCT 2914  
 Qy 897 IleThrAlaIleGlyAspValLeuGlyProSerIleAsnGlyLeuAlaSerLeuIleArg 916  
 Db 2915 GTCTCAGTTTGGAGACATATTAGGCTCTGCCATGCAAAACACAAATCTTCTCAG 2974  
 Qy 917 MetProTyrGlyCysGlyGluGlnAsnMetIleAsnPheAlaProAsnIleTyrIleLeu 936  
 Db 2975 ATGCCCTATGGCTGTGGAGACCAAAATGTGCTCTTGTCTCTTAACATCTATGATCTG 3034  
 Qy 937 AspTyrLeuThrLysLysGlnLeuThrAspAsnLeuLysGluLysAlaLeuSerPhe 956  
 Db 3035 GATATCTAAATGAACACAGACTTACTCAGAGGTCAAGTCCAAAGCCATTGGCTAT 3094  
 Qy 957 MetArgGlnGlyTyrGlnArgGluLeuLeuTyrGlnArgGluAspGlySerPheSerAla 976  
 Db 3095 CTCAACACTGGTTACAGAGACAGTTGAACAAACATGATGCTCTCAACAGCACCC 3154

Qy 977 PheGly---AsnTyrAspProSer---GlySerThrTrpLeuSerAlaPheValLeuArg 994  
 Db 3155 TTTGGGAGAGCATATGAGCAGAAACAGAGGCAACCTGGCTCAACGCTTGTTCGAAG 3214  
 Qy 995 CysPheLeuGluAlaAspProTyrIleAspIleAspGlnAsnValLeuHisAlaGlyThr 1014  
 Db 3215 ACTTTGCCCATTAAGGAGAGATGAAGATGAAGTGAACCTCTCCGCTATATCACCATC 3274  
 Qy 1015 ThrTrpLeuLysGlyIleGlnLysSerAsnGlyGluPheThrAspProGlyArgValIle 1034  
 Db 3275 ATATGGCTCTCCAGAGGCGAAGAGCAATGTGCTTTTCAAGAGCTGTGGTCACTGCTC 3334  
 Qy 1035 HisSerGluLeuGlnGlyIleAsnLysSerProValThrLeuThrAlaTyrIleValThr 1054  
 Db 3335 AACATGGCCATTAAGGAGAGATGAAGATGAAGTGAACCTCTCCGCTATATCACCATC 3394  
 Qy 1055 SerLeuLeuGlyTyrArgLysTyrGlnProAsnIleAspValGlnGluSerIleHisPhe 1074  
 Db 3395 GCCCTCTGAGATTCCTCTCACAGTCACTACCCCTGTGTCCGCAATGCCCTGTTTGC 3454  
 Qy 1075 LeuGluSer-----GluPheSerArgGlyIleSerAspAsnTyrThr 1088  
 Db 3455 CTGAGTCACTCTGGAAGACAGCAAAAGAGGAGCAATGGC---AGCCATGTATATACC 3511  
 Qy 1089 LeuAlaLeuIleThrTyrAlaLeuSerSerValGlySer---ProLysAlaLysGluAla 1107  
 Db 3512 AAAGCACTGTGGCCCTATGCTTTGGCTGGCAGAGTAAACAGCAAAAGAAAGAAAGTA 3571  
 Qy 1108 LeuAsnMetLeuThrTrpArgAlaGluGlnGlyGlyMetGlnPheTrp----- 1124  
 Db 3572 CTCAACTCATCTATGAGAGGAGTGAAGAAAGAAACAACTGTGCTTGGAGCGCCCT 3631  
 Qy 1125 ValSerSerGluSerLysLeuSerAspSerTrpGlnProArg-----SerLeuAspIle 1142  
 Db 3632 CAGAAACCAAGGACCAAGTGGGCAATTTTAAAGAACCCAGGCTCTCTGCTGAGTG 3691  
 Qy 1143 GluValAlaAlaTyrAlaLeuLeuSerHisPheLeuGlnPheGln-----ThrSer 1159  
 Db 3692 GAGATGACATCTCATGTGCTCTCGCT---TACTCAAGGCCCAAGCAGGCCCAACTCG 3748  
 Qy 1160 GluGlyIlePro-----IleMetArgTrpLeuSerArgGlnArgAsnSerLeu 1175  
 Db 3749 GAGGACCTGACCTCTCAACAACAATCGTGAAGTGAATACGAAGCAGCAAGATGCGCAG 3808  
 Qy 1176 GlyGlyPheAlaSerThrGlnAspThrThrValAlaLeuLysAlaLeuSerGluPheAla 1195  
 Db 3809 GGGGTTTCTCTCCACCCAGACAGACAGTGTGCTCTCCATGCTGTCCAATATGGA 3868  
 Qy 1196 AlaLeuMetAsnThrGluArgThrAsnIleGlnValThrValThrGlyProSerSerPro 1215  
 Db 3869 GCAGCCACATTTAC---AGGACTGGGAAGCGTCGACAGGTGACTATCCAGTTCAGGG 3925  
 Qy 1216 SerProLeuAlaValValGln----- 1222  
 Db 3926 ACAATTTCCAGCAAAATTCAGAGTGACAACAACAGCCGCTGTTACTGACAGACTGCA 3985  
 Qy 1223 -----ProMetAlaValAsnIleSerAlaAsnGlyPheGlyPheAlaIleCys 1238  
 Db 3986 TTGCCAGAGCTGTGGGAATACACATGAAAGTGAACAGGAAGAGATGTCTACCTTC 4045  
 Qy 1239 GlnLeuAsnValValTyrAsnValLysAlaSerGlySerSerArgArgArgSerIle 1258  
 Db 4046 CAGACATCTCTGAATACATAT-----CTC 4072  
 Qy 1259 GlnAsnGlnGluAlaPheAspLeuAspValAlaValLys-----GluAsnLysAsp 1275  
 Db 4073 CAGAAAGAGAAAGTTCCCTTGTGTTAGAGTCAAGCTGTGCTCAAACTTGAT 4132  
 Qy 1276 AspLeuAsn---HisValAspLeuAsnValCysThrSerPheSerGlyProGly----- 1292  
 Db 4133 GAACCCAAAGCCCAACAGCAGCTTCAAAATCTCCCTTAAGTGTGAGTTACACAGGAGCCGC 4192

QY 1293 ---ArgSerGlyMetAlaLeuMetGluValAsnLeuLeuSerGlyPheMetValProSer 1311  
 Db 4193 TCTGCTCCACATGCGATCGTGTGATGAGATGCTCTCGCTTCAATCCCTGAAG 4252  
 QY 1312 GluAlaIleSerLeu-----SerGluThrValIleValGluThrAspHisGly 1328  
 Db 4253 CCACACGTAAGAAATGCTTGAAGATCTAACATGTGAGCCGACAGAAATCCAGCAAC 4312  
 QY 1329 LysLeuAsnLeuTyrLeuAspSerValAsnGluThrGlnPheCysValAsnIleProAla 1348  
 Db 4313 CATGCTTGATTTACTTCACTGATAGGTGTCGAATCAGACACTGAGCTTGTTCTTCAAGCT 4372  
 QY 1349 ValArgAsnPheLysValSerAsnThrGlnAspAlaSerValSerIleValAspTyr 1368  
 Db 4373 CTGCAAGATGTCCTCCAGTAAGATCTGAACCAAGCCATGATGAAAGTATGATTAATCTAC 4432  
 QY 1369 GluProArgArgGluAlaValArgSerTyrAsnSer 1380  
 Db 4433 GAGACGATGAGTTCGAAATGCTGAGTCAATGCT 4468

RESULT 15  
 US-09-756-247-3  
 ; Sequence 3, Application US/09756247  
 ; Publication No. US20030180722A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Godbole, Shubhada D  
 ; APPLICANT: Boyle, Bryan J  
 ; APPLICANT: Mize, Nancy K  
 ; APPLICANT: Deng, Genhua  
 ; APPLICANT: Goodrich, Kyle  
 ; APPLICANT: Atterburn, Matthew C  
 ; APPLICANT: Tang, Y. Tom  
 ; APPLICANT: Liu, Chenghua  
 ; APPLICANT: Yeung, George  
 ; APPLICANT: Drmanac, Radoje T  
 ; TITLE OF INVENTION: METHODS AND MATERIALS RELATING TO ALPHA-2-MACROGLOBULIN-LIKE PC  
 ; FILE REFERENCE: HVS-31CIP  
 ; CURRENT APPLICATION NUMBER: US/09/756,247  
 ; CURRENT FILING DATE: 2001-01-08  
 ; PRIOR APPLICATION NUMBER: 09/649,167  
 ; PRIOR FILING DATE: 2000-08-23  
 ; PRIOR APPLICATION NUMBER: 09/540,217  
 ; PRIOR FILING DATE: 2000-03-31  
 ; PRIOR APPLICATION NUMBER: 09/684,711  
 ; PRIOR FILING DATE: 2000-10-06  
 ; PRIOR APPLICATION NUMBER: 09/560,875  
 ; PRIOR FILING DATE: 2000-04-27  
 ; PRIOR APPLICATION NUMBER: 09/496,914  
 ; PRIOR FILING DATE: 2000-02-03  
 ; NUMBER OF SEQ ID NOS: 41  
 ; SOFTWARE: Patentin version 3.0  
 ; SEQ ID NO 3  
 ; LENGTH: 4527  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 ; FEATURE:  
 ; NAME/KEY: CDS  
 ; LOCATION: (1)..(4527)  
 ; US-09-756-247-3

Alignment Scores:  
 Pred. No.: 5.56e-133 Length: 4527  
 Score: 1438.00 Matches: 428  
 Percent Similarity: 44.54% Conservative: 274  
 Best Local Similarity: 27.16% Mismatches: 557  
 Query Match: 19.57% Indels: 317  
 Db: 13 Gaps: 46

QY US-10-020-095-4 (1-1428) x US-09-756-247-3 (1-4527)  
 20 LeuAlaValAlaProGly-----ProArgPheLeuValThrAlaProGly 34

Db 28 TTGGCTTATACACGACCATTCAGAGAGAACTTCAACTCTGATGACATTAACAGCC 87  
 QY 35 IleIleArg---ProGlyGlyAsnValThrIleGlyValGluLeuLeuGlnHisCysPro 53  
 Db 88 CGGCTAAATTTCCCTC-----TCCGTCAGAAAGTTTGTGTTGAGCTGAGCCCT 135  
 QY 54 SerGluValThrValIleValGluLeuLeuLysThrIleAsnLeuThrValSerVal 73  
 Db 136 GGGTACACTGATGTTAAATTCACGGTTACTCTGAGAGACCAAGACCAAGACCAAGATTG 195  
 QY 74 LeuGluAlaGluGlyValPheGluLysGlySerPheLysThrLeuThrLeuProSerLeu 93  
 Db 196 CTAGATATCTGTGACTGAGAGAGAGGACCTTACTATATCTCTTCTTGTACACACT 255  
 QY 94 ProLeuAsnSerAlaAspGluIleTyrGluLeuLysValThrGlyArgThrGluAspGlu 113  
 Db 256 CCGTGTGTGACAGAGAAAGATGGCCAAATCCGGGTGTGGGA---GTTGAAATTAAC 312  
 QY 114 IleLeuPheSerAsnSerThrArgLeuSerPheGluThrLysArgIleSerValPheIle 133  
 Db 313 ATCAGCTTTGAGAGAGAAAGAAAGTTCTAATTCAGAGGACGGGACGACCTTTGTA 372  
 QY 134 GlnThrAspLysAlaLeuTyrLysProLysGlnGluValLysPheArgIleValThrLeu 153  
 Db 373 CAGACTGACAAACCTCTCTACACCCAGGAGAGAGATGTTTCCGCTATGTCACCATG 432  
 QY 154 PheSerAspPheLysProTyrLysThrSerLeuAsnIleLeu---IleLysAspProLys 172  
 Db 433 GATAGCAACTTCGTTCCAGTAATGACAGATCTCCAGTGGAGAACTTCAGAGATCCAAT 492  
 QY 173 SerAsnLeuIleGlnGlnThrPheLeuSerGlnIleAsnAspLeuGlyValIleSerLysThr 192  
 Db 493 AGCAACAGATTCACAGATGCTGGAAGTGTACTGACGACAGGACATTTGACCTGTCC 552  
 QY 193 PheGluLeuSerSerHisProIleLeuGlyAspTyrPheIleGlnValGlnValAsnAsp 212  
 Db 553 TTCCAACTGACACAGAGAGATGCTGGACCTACACT-----GTGGACGTGGCTGAG 606  
 QY 213 GlnThrTyrTyrGlnSerPheGlnValSerGluTyrValIleLeuProLysPheGluValThr 232  
 Db 607 GGCAGAACCTTGTGACTTTCAGTGTGAGAAATATGTGCTGCCCAAGTTTAAGGTGANA 666  
 QY 233 LeuGlnThrProLeuTyrCysSerMetAsnSerLysIleAsnGlyThrIleThrAla 252  
 Db 667 GTGGTGAACCCAGAGATTTATCAACGTGACGAGATTTTCTTAATAAAATTTGTTGT 726  
 QY 253 LysTyrThrTyrGlyLysProValLysGlyAspValThrLeuThrPheLeuProLeuSer 272  
 Db 727 AGGTACACCTATGGAAGCCCATGTGAGGACAGTGCAGATCTGTGTGT----- 777  
 QY 273 PheTyrGlyLysLysAsnIleThrLysThrPheLysIleAsnGlySerAlaAsnPhe 292  
 Db 778 -----CAGAAAGGCAAAAT-----ACTTAC 795  
 QY 293 SerPheAsnAspGluGluMetLysAsnValMetAspSerSerAsnGlyLeuSerGluLys 312  
 Db 796 TGGTATCGAGAGGTGAACGGGAACGCTTCCTGCAAAATGACAGAACCTCTCTGACAG 855  
 QY 313 LeuAsp-----LeuSerSerPro----- 318  
 Db 856 ACTGACAAACAGATGTTTCTACGACCTGTGACATGGCACCTTTGACCTCATTTGGA 915  
 QY 319 -----GlyProValGluIleLeuThrThrValThrGluSerValThrGlyIle 334  
 Db 916 TATGCTGACACCATCAATCAATTAATTTGTGCTCTGTGTGAGAGAGACAGGTGTG 975  
 QY 335 SerArgAsnValSerThrAsnValPhePheLysGlnHisAspTyrIleIleGluPhePhe 354  
 Db 976 GAGGCCAATGCCACTCAGAAATATCTACATTTTCTCCCAAAATGGATCAATGACCTTTGAA 1035  
 QY 355 AspTyrThrThrValLeuLysProSerLeuAsnPheThrAlaThrValIleValThrArg 374

Db 1036 GACACCAAGCAATTTTATCCATCCAAATTTCCCTTCAGTGGGAAGATAGATTAGAGGCC 1095  
 Qy 375 AlaAspGlyAsnGlnLeuThrLeuGluGluArgAsnValValIle----- 391  
 Db 1096 CATGATGATCTCTTCCTC-----AAGAACCATCTAGTGTTCGTGTGATT 1140  
 Qy 392 -----ThrValThrGlnArgAn----- 397  
 Db 1141 TATGACCAATGGAACCTTCACACGACCTCGTTACTGATTAACAATGAGCTAGCTCCC 1200  
 Qy 398 -----TyrThrGluTyrTrpSerGlySerAsnSerGlyAsnGlnLysMetGlnAla 414  
 Db 1201 TTATACCTTGAGACATCCGGTTGGAAATGGACACAC----- 1236  
 Qy 415 ValGlnLysIleAsnTyrThrValProGlnSerGlyThrPheLysIleGlu----- 431  
 Db 1237 -----GTTTCTCTGAGGGAAGATTTCAAATGGAAGAACTTAGTA 1275  
 Qy 432 -----PheProIleLeuGlnLysSerSerGlyLeuGlnLysAlaTyr 446  
 Db 1276 TATATCCGGAACAAGTCCAGTTACTACCAAAATGCTTACCTGACCTCCGACCTTC 1335  
 Qy 447 PheLeuGlySerLysSerSerMetAlaValHisSerLeuPheLysSerProSerLysThr 466  
 Db 1336 TACAGCAACACCCGACGCTTCCTGGCATCCACCGCTA----- 1374  
 Qy 467 TyrIleGlnLeuLysThrArgAspGluAsnIleLysValGlySerProPheGlnLeuVal 486  
 Db 1375 -----AACGGCCCTTGAATGTGGCCAGCCGAGCCGAGGAAGTGTCTG 1413  
 Qy 487 Val-----SerGlyAsnLysArgLysLeuLysSer 497  
 Db 1414 GTGATATTATCATGACCCGCGGATGACAGCCCTGACCAAGATC--AGCTTCTCC 1470  
 Qy 498 TyrMetValValSerArgGlyGlnLeuValAlaValGlyLysGlnAn----- 513  
 Db 1471 TACTATTATTAAGGAAGAGTTGTGTATGAGAGGGGCGAACAACATCGAACTTAG 1530  
 Qy 514 -----SerThrMetPheSerLeuThrProGluAsnSerTrpThrPro 527  
 Db 1531 AAGAAAGCATGAAAGCT 1590  
 Qy 528 LysAlaCysValIleValTyrTyrIleGlnAspAspGlyGlnIleIleSerAspValLeu 547  
 Db 1591 GATCCTTCCTGATCTATGATGATCTTTTCCAGTGAAGGTGTGTAGTCACAAATT 1650  
 Qy 548 LysIleProValGlnLeuValPheLysAsnLysIleLysLeuTyrTrpSerLysValLys 567  
 Db 1651 CAGTTCTCAGTGGGATGTGCTTGCACATCAGGTTTCCCTTGCCTTCTCCCTCCAG 1710  
 Qy 568 AlaGluProSerGlyLysValSerLeuArgIleSerValThrGlnProAspSerIleVal 587  
 Db 1711 CAGCTTCAGAGAGCAAGATGAGTGCAGCTGCAGCAGCT--CCCGATCCCTGTGT 1767  
 Qy 588 GlyIleValAlaValAspLysSerValAsnLeuMetAsnLysSerAsnAspIleThrMet 607  
 Db 1768 GCGCTCCGGGGGTGATGATGATGTCTTACTGCTTAGGCCA----- 1809  
 Qy 608 GlnAsnValValHisGluLeuGlnLeuTyrAsnThrGlyTyrTyrLeuGlyMetPhe--- 626  
 Db 1810 -----GACAGAGAGCTGAGCAACGCTCTGTCTAT--GGGATGTTTCCA 1851  
 Qy 626 ----- 626  
 Db 1852 TTCTGTATGTCTACTACCTTATCAAGTGCAGTATGATCAGTGTCCAGTGTCTGGC 1911  
 Qy 626 ----- 626  
 Db 1912 CCATGGGACTTTCTCAGGCCCTCATTTGACCAATGCCCAAGGGCATGAGCCAGCT 1971  
 Qy 627 -----MetAsnSerPheAlaValPheGlnGlu 635  
 Db 1972 TCATATTCTGAGGCGCTCTCTCTGTAAGCAGCAGCATTTTCAGCTTTTCCGGGAGC 2031

Qy 636 CysGlyLeuTrpValLeuThrAspAlaAsnLeuThrLysAspTyrIleAspGlyValTyr 655  
 Db 2032 GTGGGCGTGAATAATGATGTCATATGCAAAATCAAGAGCCA--GTAGTTGTCAGTAC 2088  
 Qy 656 AspAsnAlaGluTyrAla-----GlnArgPheMetGlu 666  
 Db 2089 AGATCTCCAAATACAGCACTGCTATGGTGGGGGTGTCATCAAGAGGCTTTTGAAGTCA 2148  
 Qy 667 GlnAsnGluGlyHisIleValAlaPheIleAspPheSerLeuGlySerSerProHisVal 686  
 Db 2149 TCATCTCTTATCATCAAGACAG-----GATTTCTCAGTTC 2184  
 Qy 687 ArgLysHisPheProGluTrpIleTyrPheAspThrAsnMetGlyTyrArgIleTyr 706  
 Db 2185 CGCAGACTTCCAGAGACTGCTGCTGGATCTGTTCTTATGTGTAATCTGGGGAG 2244  
 Qy 707 GlnGluPheGluValThrValProAspSerIleThrSerTrpValAlaThrGlyPheVal 726  
 Db 2245 GAGCGGTCCACGTCAAGTTCCTGACCGCATCACCGAGTGAAGCGAGTGTCTGC 2304  
 Qy 727 IleSerGluAspLeuGlyLeuThrThrThrProValGluLeuGlnAlaPheGln 746  
 Db 2305 ACTTCCAGTCAAGAGCTTGGGCTTTCACCCACT--GTTGACCTAATCTGCTTTCAG 2361  
 Qy 747 ProPhePheIlePheLeuAsnLeuProTyrSerValIleArgGlyGluGluPheAlaLeu 766  
 Db 2362 CGTTCCTTTGTACCTGACATCTCCCTTACTGATGTCCGTGGGAATCTTTCGCTT 2421  
 Qy 767 GlnIleThrIlePheAsnTyrLeuLysAspAlaThrGluValLysValIleIleGluLys 786  
 Db 2422 ACTGCCACCATCTTCAATTACTTAAGAGATTCATCAGGCTTACAGCTGCGGTAA 2481  
 Qy 787 SerAspLysPheAspIleLeuMetThrSerSerGluIleAsnAlaThrGlyHisGln 806  
 Db 2482 TCGCATGATGACGACGATGATATGATGCGGAGATTCT--CAGACC 2523  
 Qy 807 ThrLeuLeuValProSerGluAspGlyAlaThrValLeuPheProIleArgProThrHis 826  
 Db 2524 TCCAGTTGTCTGTGTGTATGACGCAAAACCAACATCGAATCAGCTGTCAAA 2583  
 Qy 827 LeuGlyGluIleProIleThrValThrAlaLeu----- 837  
 Db 2584 TTGGGTCACTTAATCTTACTTACTATGACAAAGATTCTGCAGCAGATGACATGTGG 2643  
 Qy 838 -----SerProThrAlaSerAspAlaValThrGlnMetIleLeu 850  
 Db 2644 GGCAGAGAGGGTTTGTCCCAAAAGGCCGGAAGTACCGCTCATCAAGCCAGTCTC 2703  
 Qy 851 ValLysAlaGluGly-----IleGluLysSerTyrSerGlnSerIleLeuLeuAspLeu 868  
 Db 2704 GTCAACCTGAGGAGATCGTGTGAGAGACACAGC-----TCATTG 2748  
 Qy 869 ThrAspAsnArgLeuGlnSerThrLeuLysThrLeuSerPheSerPheProAsnThr 888  
 Db 2749 CTGTGCCAAAGAAAGAGGAGCATGCTGATGTCTCCCTGAGGCTCCACAGTGCATTT 2808  
 Qy 889 ValThrGlySerGluArgValGlnIleThrAlaIleGly----- 901  
 Db 2809 GTTCTGACTCGACCAAGGCTTATGTACGTTCTGTGGTAAAGCATTTAGATTCTTGAC 2868  
 Qy 902 -----AspValLeuGlyProSer 907  
 Db 2869 TCAGAAAGAAAGCGAATGAGGCTGCAGAGGTGTGAGAGACATTTAGGCACAGCC 2928  
 Qy 908 IleAsnGlyLeuAlaSerLeuIleArgMetProTyrGlyCysGlyGluGlnAsnMetIle 927  
 Db 2929 CTCAGAACTGTGATGTGTGTGCAAGATGCCAGTGGCTGTGGCAGACGAATGTGTC 2988  
 Qy 928 AsnPheAlaProAsnIleTyrIleLeuAspTyrLeuThrLysLysGlnLeuThrAsp 947  
 Db 2989 TTGTGTCTCCCATCATATATGTCTTTCAGTACTGAGAAAGCAGGCTGTGACGAG 3048



QY 948 AsnLeuLyGluValAlaLeuSerPheMetArgGlnGlyTyrGlnArgGluLeuLeuTyr 967  
 Db 3049 GAGATCAAGTCTCGGACAGTGGTTCCTGGAATAGGGTACCAAGAGCTGAGATTC 3108  
 QY 968 GlnArgGluAspGlySerPheSerAlaPheGlyAsnTyrAspProSerGlySerThrThr 987  
 Db 3109 AAAACACAGAAATGGCTCATACAGTGGCTTTGGGAGCCGAGATGGAAATGGAAACACATGG 3168  
 QY 988 LeuSerAlaPheValLeuArgCysPheLeuGluAlaAspProTyrIleAspIleAspGln 1007  
 Db 3169 CTGACAGCGCTTGTCTCAAAATGCTTGGCCAAAGCTCAGAAATTCATCTTCATTTGATGCC 3228  
 QY 1008 AsnValLeuHisArgThrTyrThrTyrLeuLyGlyHisGlnLySerAsnGlyGluPhe 1027  
 Db 3229 AAGAACCTCAGAGATGCTCTCAAGTGGAGGAGAAACAGGTCCCGCTGCTCTAT 3288  
 QY 1028 TrpAspProGlyArgValIleHisSerGluLeuGlnGlyGlyAsnLySerProValThr 1047  
 Db 3289 GCCAACGTGGAAATCTCTTCACACAGCTATGAGGGTGTGATGATGATGAGGCTCTCC 3348  
 QY 1048 LeuThrAlaTyrIleValThrSerLeuLeuGlyTyrArgLySerGlnProAsnIleAsp 1067  
 Db 3349 TTGACGTGGTATGTCAACAGTCGATTGCTGGAGATGGGAAAGATGATGATGACCAACG 3408  
 QY 1068 ValGlnGluSerIleHisPheLeuGluSerGluPheSerArgGlyIleSerAspAsnTyr 1087  
 Db 3409 GTGAGTCAGGCTCTACGAGTGTCTCAAGAAATTCG---GCCACTCCACGACCAACTCTAC 3465  
 QY 1088 ThrLeuAlaLeuIleThrTyrAlaLeuSerSerValGlySerProLyAlaLyGlu--- 1106  
 Db 3466 ACACAGGCGCTGTGGCTTACATTTTCTCTGGCTGGGAAATGACATCAGAAACATT 3525  
 QY 1107 AlaLeuAsnMetLeuThrTrpArgAlaGlnGlnGlyGlyMetGlnPheTrpValSer 1126  
 Db 3526 CTCCTTAAACAGTATGATCAACAGGCTATCATCTCAGAGAAATCCATTACTGAGCCAG 3585  
 QY 1127 SerGluSerLyLeuSerAspSer---Trp---GlnProArgSerLeuAspIleGlu 1143  
 Db 3586 AAACCTACTCCATCATCGAAACCCAGCCCTGTGCTGAGCTGGCGCTGATGATGAGAA 3645  
 QY 1144 ValAlaAlaTyrAlaLeuLeuSerHis-----PheLeuGlnPheGlnThr 1158  
 Db 3646 CTCACAGCATATGATGCTTGGCCAGCTTACCAACCCAGCTGACTCAAAAAGAGATA 3705  
 QY 1159 SerGluGlyIleProIleMetArgTrpLeuSerArgGlnArgAsnSerLeuGlyGlyPhe 1178  
 Db 3706 GCGAAGCCCACTAGCATATGCTGCTGGTGGCCAAACACATCATATGGGGCTTC 3765  
 QY 1179 AlaSerThrGlnAspThrThrValAlaLeuLyAlaLeuSerGluPheAla-----Ala 1196  
 Db 3766 TCTTCTACTCAGGATATCTAGTGTGCTCTCAAGCTCTTGCCAAATATGCCACTACCGCC 3825  
 QY 1197 LeuMetAsnThrGluArgThrAsnIleGlnValThrValThr----- 1210  
 Db 3826 TACATGCCATCTGAGGAGATCAACCTGTTGTAATAATCCATGAGAATTTCCAGCGCACA 3885  
 QY 1211 ---GlyProSerSerProLeuAlaValValGln-----Pro 1223  
 Db 3886 TTCACATATCAGTCACTTAAACAGATTTGTAATTCACAGAGATACCTGCCCAATGTCTCT 3945  
 QY 1224 MetAlaValAsnIleSerAlaAsnGlyPheGlyPheAlaIleCysGlnLeuAsnValVal 1243  
 Db 3946 GGAATGTACACGTGGAGGCTTCAGGCCAGGGCTGTGTATGTGCAGACGGTGTGGAGA 4005  
 QY 1244 TyrAsnValIleAlaSerGlySerSerArgArgArgSerIleGlnAsnGlnGluAla 1263  
 Db 4006 TACAATATCTCCCTCC-----ACAAATATGAAGACC 4038  
 QY 1264 PheAspLeuAspValAlaValLyS-----GluAsnLySAspAspLeuAsnHis 1279  
 Db 4039 TTTAGCTTATGATGGAATAGAAAAGCTAGATGTGACCAACCGACTTACCTCGATCC 4098  
 QY 1280 ValAspLeuAsnValCysThrSerPheSerGlyPro---GlyArgSerGlyMetAlaLeu 1298

Db 4099 TTGACTCTCACTATTCAACACGATTATGTGGGAGCCGTAGCTCTTCAATATGGCTATT 4158  
 QY 1299 MetGluValAsnLeuLeuSerGlyPhe-----MetValProSerGluAlaIleSerLeu 1316  
 Db 4159 GTGGAAGTGAAGATCTATCTGGGTTCACTAGTCCCAATGAGGGAGCAACATCAGTACTTCTC 4218  
 QY 1317 SerGlu---ThrValIleLyValGluTyrAspHisGlyLyLeuAsnLeuTyrLeuAsp 1335  
 Db 4219 CAGCAACCCCTGGTGAAAGAGTTGAATTTGGAACTGACACACTTAACATTTACTTGAT 4278  
 QY 1336 SerValAsnGluThrGlnPheCysValAsnIleProAlaValArgAsnPheLyValSer 1355  
 Db 4279 GAGCTCATTAAGACACTCAGACTTACCTTCAACATCAAGCAAGTGTGCTGCTCACCC 4338  
 QY 1356 AsnThrGlnAspAlaSerValSerIleValAspTyrTyrGluProArgArgGlnAlaVal 1375  
 Db 4339 AACTTGAAACACGACCACTCAAGGCTATATGACTATCTACCA----- 4383  
 QY 1376 ArgSerTyrAsnSerGluValIleLeuSer-SerCysAspLeuCys 1390  
 Db 4384 -----GTTCTTTAAATTATCTCAGTACCAATTGTGT 4417

Search completed: January 17, 2004, 08:49:35  
 Job time : 1238 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 15, 2004, 18:10:30 ; Search time 28 Seconds  
(without alignments)  
4904.603 Million cell updates/sec

Title: US-10-020-095-4

Perfect score: 7348

Sequence: 1 MGGPPLTAHLVCVTAL.....HSSVIFFCFLYFMEIWL 1428

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-Processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :  
1: pir1:\*  
2: pir2:\*  
3: pir3:\*  
4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1854.5	25.2	1508	2 T27828	hypothetical prote
2	1850	25.2	1519	2 T27829	hypothetical prote
3	1638	22.3	1507	2 T18544	alpha-2-macroglobu
4	1451	20.3	1476	2 JCS143	alpha-macroglobuli
5	1455	19.8	1477	2 A29952	alpha-1 proteinase
6	1440	19.6	1474	1 MAHU	alpha-2-macroglobu
7	1424	19.4	1503	2 T43166	alpha-2-macroglobu
8	1417	19.3	1487	2 S15904	alpha-1 proteinase
9	1412.5	19.2	1472	2 A26122	alpha-2-macroglobu
10	1402	19.1	1451	2 B41185	alpha-2-macroglobu
11	1396.5	19.0	1482	2 S13495	pregnancy zone pro
12	1395.5	19.0	1476	2 A41185	alpha-2-macroglobu
13	1395	19.0	1473	2 A20872	ovostatin precursor
14	1384	18.8	1500	2 A42210	alpha-1-macroglobu
15	1374.5	18.7	1464	2 JCS144	murinoglobulin pre
16	1373.5	18.7	1495	2 S27001	alpha-2-macroglobu
17	1373.5	18.0	1699	2 T14074	complement compo
18	1373.5	18.0	1699	2 T30885	complement compo
19	1373.5	18.0	1699	2 T30885	complement compo
20	1373.5	18.0	1699	2 T30885	complement compo
21	1373.5	18.0	1699	2 T30885	complement compo
22	1373.5	18.0	1699	2 T30885	complement compo
23	1373.5	18.0	1699	2 T30885	complement compo
24	1373.5	18.0	1699	2 T30885	complement compo
25	1373.5	18.0	1699	2 T30885	complement compo
26	1373.5	18.0	1699	2 T30885	complement compo
27	1373.5	18.0	1699	2 T30885	complement compo
28	1373.5	18.0	1699	2 T30885	complement compo
29	1373.5	18.0	1699	2 T30885	complement compo

30	934	12.7	1663	1 C3HU	complement C3 prec
31	921	12.5	1676	1 C5HU	complement C5 prec
32	905	12.3	1680	1 C5HU	complement C5 prec
33	904	12.3	1666	1 C3GP	complement C3 prec
34	894	12.2	1642	2 S21018	cobra venom factor
35	857.5	11.7	1620	2 S21045	complement protein
36	736.5	10.0	609	2 A41081	alpha-1-inhibitor
37	694	9.4	1230	2 T30517	complement C4 - Q2
38	688.5	9.4	1230	2 T30517	complement C4 - Q2
39	467.5	6.4	1536	2 E72310	hypothetical prote
40	432.5	5.9	726	2 A27602	complement C3 - ra
41	290.5	4.0	1892	2 C97804	hypothetical prote
42	287	3.9	1653	2 B91052	hypothetical prote
43	287	3.9	1653	2 F85896	hypothetical prote
44	283	3.9	1653	2 G65028	hypothetical prote
45	269.5	3.7	1644	2 AC0823	probable lipoprote

## ALIGNMENTS

## RESULT 1

T27828

hypothetical protein ZK337.1a - Caenorhabditis elegans

C/Species: Caenorhabditis elegans

C/Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 17-Mar-2000

C/Accession: T27828

R/White, S.

submitted to the EMBL Data Library, November 1996

A/Reference number: Z20426

A/Accession: T27828

A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: DNA

A/Residues: 1-1508 <MIL>

A/Cross-References: EMBL:Z82090; PIDN:CA80506.1; GSPDB:GN00019; CESP:ZK337.1a

A/Experimental source: clone ZK337

A/Genetics:

A/Gen: CESP:ZK337.1a

A/Map position: 1

A/Intons: 36/2; 116/3; 201/1; 267/2; 338/3; 413/3; 527/3; 556/3; 691/3; 714/1; 760/2; 8

C/Superfamily: alpha-2-macroglobulin

Query Match 25.2% Score 1854.5; DB 2; Length 1508;

Best Local Similarity 31.8%; Pred. No. 6.3e-97;

Matches 476; Conservative 292; Mismatches 562; Indels 165; Gaps 44;

QY	15	VCTAALAVAGPRFLVTAAGIRPGANTYIGVLLHCPQVTVKAEILKTASNTLSVTL	74
DB	28	VSTTAAPVKA-TYMLVAPAVVRPDQPSVCNMLKQATDEDMIVRIEVRTERNETIAR	86
QY	75	EAEGVFEKSGFKTLPLPLNSAD-EIYELRVGRQD-EILFNSNTRLSFEKRTISV	131
DB	87	VISNN-KPGIAQTVSLSEMPASQSLTPROSKYIRGETLMAELIFENENELKYDOKALSV	145
QY	132	FIQTQKALYKPKQEVKRVLTLSFDPKPKYTSNLTIIDPKSNLTQOMLSQSDGVTISK	191
DB	146	FIQTBRALYRPSLVRYRAIVKSKDKRYGNATIKIDPSPNLSQITIGVTLDGVSIG	205
QY	192	FTQLSHPLIGDWSIQVQ---VNDQYYQSPQSEVYLPRKEVTLQTPLYCSMNSKHLN	247
DB	206	ELQALBETLLGWMFLFEVERISNGVQDS--STVDTYVLPRKEVNIKSSFTIND-DLS	261
QY	248	GTTTAKTYTGKGVKQDVLTFL-LPLSPF-----GKKK---ITKTFKIN--GS	289
DB	262	VFDVAKTYTGKGVKQDVLTFL-LPLSPF-----GKKK---ITKTFKIN--GS	321
QY	290	ANFSFNDSEKMKVMDSSNGLSEYLDSSPGVEILTYTESVTGSRNVSTVFVKOHY	349
DB	322	AAVFSNDELK-----KLLHEMGGSIRIYASTEDTELEARNATHOISTFREEV	373
QY	350	IIFFDYTYTLKPSINFTATYKVRADGN--QTLLEBRNNVITVTOANTYEVSGNS	407
DB	374	KLDVERKQDTPFKRGTLVNVVALKQMDTPVYKATLPKR---VQSTFYNY--YNHDS	427

```

Qy 408 GNOQMEAVOKINTYVPOSGTFKIEPILEDSELOKAYFLGSKSMVAHSLFK----- 461
Db 428 SLOGEKRTKIYEVDAHGTSLVLTQPIINCTSARIEAH-YDIGKDNFATPIYSLSLYEA 486
Qy 462 --SPSKTYIOUKTRDEN-IKVGSPFELVVSNGKRLKELSYVVSROQVY-----AVGKON 513
Db 487 AVSPKTSFLQILADNEGADVGSLSFSLKATOPSLITYOVMSRSNIYVSQOMTVNSEH 546
Qy 514 STMFSLTPENSWTPKACVIYV-YIEDDEGIISDVLEKIPVOLVFNKIKLYMSKYKAPSE 572
Db 547 ATI-SFPATANMAKSKRLIVYAIIESQEVLDALDFVBEQIFQONVALSIDKQAVEBQ 605
Qy 573 KVSRLISVTPDSIVGIVADVSKVNLMAASNDITMENVHLELYNTGYLYGMF----- 626
Db 606 NVKFKVT-SDKNSFVGLLVVDQSVLLKTKGNDITREKVEQOLEMYDSDNNVGGGFGPRPW 664
Qy 627 -----MNSFAVQEGCLMVLTDANLTD-----YITGVYDNA 658
Db 665 EAIIDRRKRSIWRPMWIGIGSDAOSIFSNAGLVLTDLALYREPOREFMSVMMMDGADGMA 724
Qy 659 EYARFEMENEGHIVDHFSLGSS---PHYRKHPPEPMWLTDTNMGRIRIQEEVTVPD 715
Db 725 EAA--FAAPMG-----GSSPPPTVYKFPFHTWMSDLN-STGGEVMEIAPD 771
Qy 716 SITSMVATGFVISEDGLGLTTTPEVLOAFOPFELFNLPSYVIRGEEFALEITIFNYLK 775
Db 772 TITSMVASTFAINENGAVAPTTSKLVRFPFQIQLNLYAVARGEFALVAVFYNYME 831
Qy 776 DATEVKYII--EKSDKPIILMTSSEI--NATHQOQTLVSEBQAT--VLPFIRPTH 827
Db 832 KEQOVTVYLYKDYKOSGDLLKKDGTIVARDEVGOONRIVSVAGGTSKAYVFPVPSI 891
Qy 828 GEIPITVYALPTASDAVTOMILYKABGIEKSYQSILDLTDNRLOSTLTKTSFSPSPN 887
Db 892 GEIPIHSAISOGSDAVENMLRDPQCYKVDKRNIPFIIDLNNSDPS-KNLELIPND 950
Qy 888 TVTSEKRVQITAGVGLPSINGLASLIRMEYGGCEOMNINFAPNIYILDYTKKQULTD 947
Db 951 VVDSQKARLVDIGMGMGPVNLNAHKLVQMPYGGEGQMLNLVNLIVVKTLRATNNES 1010
Qy 948 NLKELASFPKQYQRELLYQREDSFSAFGNYPGSGTWLSAVVLCFLEADPYIDIDQ 1007
Db 1011 QLEFKAIKFIQOGLQRELLYKRADNSFSFADSDKAGSTWLTAFVVSFHHAKYAFVDP 1070
Qy 1008 NVLHRTYTWLKGHKSNGEFPMDPGRVHSELQGG-NKSPVLTAYITSLIGYKXQPN 1066
Db 1071 NVISRAVAFNLNSQOMESGAFERGBVHHKMDQGAQDGGVALTAFVLISIL-----ENG 1125
Qy 1067 DVQESIHLESEFSGISDNTYLLALITYALSVSFPKAKELANMLTWK--AEDEGQOFW 1124
Db 1126 ENGRAVYTLLEHGLDEVSGNATYMAVVALOLAKSKQAKAFENLKHKKIYKSGDVKA 1185
Qy 1125 VSSE--SKLSDS---WQPSLDIEVAAYALLSHFLOFQISEGIPIMKWLSRQNSIGGF 1178
Db 1186 SAQKVEKELAKESRAYMFOARPDVDETTSYAVLSYLAQNOTSESLIIRMLVVSQNEIGGF 1245
Qy 1179 ASTODTTALKALSEBALMTERTNIQYTV-TGPSSPS-----PLAVVQPM----- 1224
Db 1246 TSTDYTWALQALSSYAAVYTSDDHTSQVITLKNKHTHSFDINIRNAIVLOSQYOLSLND 1305
Qy 1225 AVNISANGFPAICQLNVVNVVVKASGSSRRRSIONQEPDLDAVAENKODLHNHVLND 1284
Db 1306 AVSINANGTVVFKQLSYST-----YDSDLANDAPFCSEIETIRAG-NRLQDL 1355
Qy 1285 CTSGSGPGRSGMALMEVNLGFWVPSEALSSETVK-----KVEYDHGKMLVLDSEVNE 1339
Db 1356 CCNTRYRCKSNMALAEIDALSGYRFDQVHTLTSIEDLOKVEKODTKKNVYFNLGG 1415
Qy 1340 TQFCVNIIPAYANFKYSNTQDASVSIUDYEPBRQAVNSYNSVGLSCDCLSDVQGRP 1398
Db 1416 RPYCLSTYSVDVTVADOKPANFRLVYDPEEQIKMTYAKQTRSLQEKGC--EDCW 1472

```

```

RESULT 2
T27829
hypochemical protein ZK337.1b - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 17-Mar-2000
C:Accession: T27829
R:White, S.
submitted to the EMBL Data Library, November 1996
A:Reference number: Z20426
A:Accession: T27829
A:Status: preliminary; translated from GB/EMBL/DBD3
A:Molecule type: DNA
A:Residues: 1-1519 <MIL>
A:Cross-references: EMBL:Z282090; PIDN:CAB05007.1; GSPDB:GN00019; CESP:ZK337.1b
A:Experimental source: clone ZK337
C:Genetics:
A:Gene: CESP:ZK337.1b
A:Map position: 1
A:Insertions: 36/2; 116/3; 201/1; 267/2; 338/3; 413/3; 527/3; 556/3; 691/3; 714/1; 771/2; 84
C:Superfamily: alpha-2-macroglobulin

Query Match      25.2%   Score 1850; DB 2; Length 1519;
Best Local Similarity 31.6%; Pred. No. 1,1e-96;
Matches 477; Conservative 290; Mismatches 563; Indels 180; Gaps 44;

Qy 15 VCTAALAVAGPREFLVAPGIIIRPGGNVTIGVELHEHCFSQVTKAELKTASNLTVSVL 74
Db 28 VSTTAAPVKA-TYMLVAPAVAPRDPQFVCMNLQATDEDMIVALEVTEFNEETIAAR 86
Qy 75 EABVEPKSGFKTLPLSLPLNSAD--EYELRVGTQD-EILFNSRSLSEFYRISV 131
Db 87 VISNL-KPGIAQVLSLEMPAQSLTPROGYKLYIREETINAEILFENNELKQDQALSV 145
Qy 132 FIOTDRLYKPKQKQKFRIVTLFSDSKPYKTSINILIKPKSKYLQOMLSQSDLGVISK 121
Db 146 FIOTDRLYKPAVLVRAIVKSDIKPYGNATIKIPPSRNLISQITGVTLDRGVSG 205
Qy 192 TPOLSHPILGDSIQVQ-----VNDQTYVQFQVSEYVLPKFEVTLQTPLYCSMNSKHLN 247
Db 206 ELQABETLGLDFIEVENSNGVQDS---SFTVDYVLPKFEVNIKTSSFITIND-DLS 261
Qy 248 GTTATKYTKPKYKGVNTLTF-LPLSPW-----GKKKN---ITKTKIN--GS 289
Db 262 VFVDAKTYGKGAAGAKVSLLEPMHRHAMVPTIIDENGVKKEELMVERIVKLNROGE 321
Qy 290 ANFSFNDDEKKNVMDSSNGSEYLDLSSPGPVILTTVYESVYIGISRNSTNPFQOHY 349
Db 322 AAVVFSNDELK-----HKLHEWGGSIIRIVASVETDIEIRNATHQISTREEV 373
Qy 350 IIEFPDYTVLKPSLNFTATVKTADGN--QLTEERRNNVITYQRYTEYMGSGNS 407
Db 374 KLVVEKQDIFKGLLYNVVVALKQMDDPVKATLTKR-----VOVTFNYNP--YHHD 427
Qy 408 GNOQMEAVOKINTYVPOSGTFKIEPILEDSELOKAYFLGSKSMVAHSLFK----- 461
Db 428 SLOGEKRTKIYEVDAHGTSLVLTQPIINCTSARIEAH-YDIGKDNFATPIYSLSLYEA 486
Qy 462 --SPSKTYIOUKTRDEN-IKVGSPFELVVSNGKRLKELSYVVSROQVY-----AVGKON 513
Db 487 AVSPKTSFLQILADNEGADVGSLSFSLKATOPSLITYOVMSRSNIYVSQOMTVNSEH 546
Qy 514 STMFSLTPENSWTPKACVIYV-YIEDDEGIISDVLEKIPVOLVFNKIKLYMSKYKAPSE 572
Db 547 ATI-SFPATANMAKSKRLIVYAIIESQEVLDALDFVBEQIFQONVALSIDKQAVEBQ 605
Qy 573 KVSRLISVTPDSIVGIVADVSKVNLMAASNDITMENVHLELYNTGYLYGMF----- 626
Db 606 NVKFKVT-SDKNSFVGLLVVDQSVLLKTKGNDITREKVEQOLEMYDSDNNVGGGFGPRPW 664
Qy 627 -----MNSFAVQEGCLMVLTDANLTD-----NLT 647
Db 665 EAIIDRRKRSIWRPMWIGIGSDAOSIFSNAGLVLTDLALYREPOREFMSERRLNTPGLT 724

```

```

Qy 648 KDIDGVYDNAMEAEERFMEENEGHIVDIDHDSIGSS---PHYRKHPETWIDTMNGYR 704
Db 725 VMMMDGAPGMAEAA--FAAPRMG-----GSSPPPTVRKPFPHIMWSDLN-SNS 771
Qy 705 IYQEEVYVPSDITSVATGATGPISEDLGILTTTPVELAQFQPFPLNLPYSVIRGEFP 764
Db 772 GVEVEIEAPDITTSVASTFAINENEGIGVAPITTSKLVFRPPFQLMLPYAVVRGGEKF 831
Qy 765 ALEITFNVLKATEKVKII--EKSDKFDILMTSSBI---NATGHOQTLTVSEDAQT-- 817
Db 832 ALLVIVFNMKEKQDVTYTLKIDKDSGYDLKKDGVVARDVGGQNRIVSVAAGGTSK 891
Qy 818 -VLPFIRPTLGEIPITVTALSPASDAVTOMILVKAEGIEKSYSGSILLDLTDNRLOST 876
Db 892 AVYFPFIVPSIGEIPVHISAIASQGGDAVENMLRVDPQGVKVDRIPIYIDLNNSDPS 951
Qy 877 LKTSFSPFPNTVTSSEVQITAGIDVLCPSINGLASLIRMPYGCSEQMINIPAPITYL 936
Db 952 -KNEELIWPNDVDVDSQKARLDVIGPMGPRVILNNAKLVQMPYGCCEQNMMLVPIVILV 1010
Qy 937 DVLTKKQOLTDMLKEKALSFMROGYORELLYOREDSGSAFGVDPDSSTWLSAFLRGE 996
Db 1011 KYLRATNRESQLETKAIKFIIEGIGRELTYKADNSFSAFGSDKASFTMLAFVVRSGF 1070
Qy 997 LEADPYIDIDQNVLTHTYTLWLGKHOKSGEFMDPGRVISELQGG-NKSPVTLVAVTYS 1055
Db 1071 HIAKQYAFVDPNVITSAVAFVINSQOMESGAFARGEVHHKMDQGGQDQGVALTAFVLIS 1130
Qy 1056 LIGYRKYQPNIDVQESIHFLSEFSRGISDNTLTALITYALSSVSGPRAKEALNMLTWR- 1114
Db 1131 IL-----ENGEMNGKAVYLEKHLDSEVSGNAAYTMAVAYALQAKSKQAKAFENLKKKK 1185
Qy 1115 -AEQSGMGFWVSSE--SLSDS---MQPRSLDIEVAVALLSHRLQOTSEGIPIMW 1167
Db 1186 IVEKSGDVAFSAQKVEKLKESRAYMFOARPDLETTYSYAVLSYLAQNTSESLTIKW 1245
Qy 1168 ISROBNSLGGFASTODTVAALKALSEFALMTERNTIOVTV-TGPSSS-----PLAV 1220
Db 1246 LVSQNELGCFSTQDTWALQALSYAVTSDKHTSQVTLINGHTHSFDINTIRNALV 1305
Qy 1221 VQPM-----AVNISANGFPAICQLNVVNVKVASGSSRRRRSICNOQAFDLDAVAKEN 1273
Db 1306 LQSYQLSLINDAVSINANGTVFAQLSYSY-----YRDSLNDAPFPCSGEIKEI 1356
Qy 1274 KODLHVLDLVCTSPSGPERSGMALEBVULLSGFMPSPSAISSEYVK-----KAEYDGG 1328
Db 1357 RAG-NRLQDLCCNTRPKSKNMALAEIDALSGYRDAEQVHTLTISIEDLQREVEKEDJT 1415
Qy 1329 KLNLYDSVNETQFCVNI PAVNRFKYSNTQDASVSIIDVYEPBRQAVRSYNSFVKLSGD 1388
Db 1416 KNNVFNPLGGRPVCLISLSDVTYQVADQKPNFRLVYDYPDEQLKMTYAAKQTRSLQE 1475
Qy 1389 LCSDVQGCPR 1398
Db 1476 KCG--EDCWP 1483

```

## RESULT 3

```

alpha-2-macroglobulin - Limulus sp.
C:Species: Limulus sp.
C:Date: 02-Sep-2000 #sequence_revision 02-Sep-2000 #text_change 15-Sep-2000
C:Accession: T18544
Ritwaki, D.; Kawabata, S.; Miura, Y.; Kato, A.; Armstrong, P.B.; Quigley, J.P.; Nielsen,
Eur. J. Biochem. 242, 822-831, 1996
A:Title: Molecular cloning of Limulus alpha-2-macroglobulin.
A:Reference number: Z18961; MUID:97175055; PMID:9022715
A:Accession: T18544
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-1507 <IMA>
A:Cross-references: EMBL:D83196; NID:d1096438; PID:d1020631; PIDN:BAAL9844.1

```

A:Experimental source: adult  
C:Superfamily: alpha-2-macroglobulin

Query Match 22.3%; Score 1638; DB 2; Length 1507;  
Best Local Similarity 28.9%; Pred No. 1,3e-84;  
Matches 449; Conservative 270; Mismatches 559; Indels 278; Gaps 43;

```

Qy 28 FLVAPGIIRPGNVITIGVELLEHCPQVTVKALLKTASMLTVS-----VLEAG 78
Db 29 FILTAPKSLTPKSNILMLHFD-----IKINGFLRIGVKQDDGNVAETEV 76
Qy 79 VFEK--GSFKTLTPS-----LPNSADEIYELRVGTODELIFNSSTRLSFETKIS 130
Db 77 SFNKDNPSSSIQLTIPSGEVVRPFLYANGSY---SSPSNDPFEEKD--INMKDKLI 130
Qy 131 VFIQTDKALYKQKQVFRITVTLFSDFK--PKTSNLTILQPKSNLQOMLSQOSDLCV 188
Db 131 VVVQTDKPLKPGQVVKRILPTTPDLKLVKETTIGSFQIENPDGIVLGYPMLSPACGI 190
Qy 189 ISKTPQLSSHPIIGDMSIQOVNDQTYQSPQVSEYVLPKEVTLQTPLYCSMNSKHLNG 248
Db 191 AQFELALDDEPTGYKMRILKGNIEDTEIENFEVKEVYLPKEVKTLPSSYLLTNADSLTW 250
Qy 249 TTTAKTYGKPYKGVTLTFLPLSFWKKNITKTFK-----INSANFSFND 297
Db 251 KICAOYTGQPFEGTFVAETNVVKKYMKMEKEGVPVHKSGLDIGCLDVTVNSSA-LGFNEQ 309
Qy 296 EKMVMDSNGLSYELDSSPGVEILTLYTESVTVIGSNVSTVFPQHOVYIEPDT 357
Db 310 RL-----STRAVNMFABVIEKGGIKKQKADNTSITRTSNPLNIMYLEPT 352
Qy 358 T--VLKESLNFATVKTVRADGNQLTEERRNNVITVQRYNEYVSGNSGNOKMA 414
Db 353 SGKYLKGLKLPYGGKLTKEKPEGTAPGEQ-----IELCRADERNNRKKMLEKIRA 406
Qy 415 VQKINYTPQSGTFKIEF-PILEDSSQLKAYFL-----GSYSMAVHSL---FKS 462
Db 407 CKE--FTSDEAGIIFKTVPPQPTDITSPFRKAKALQYCKQKODNKLNGQHSFTVSSWS 464
Qy 463 PSKTYIQLKTRBNKVSFPFL---VVSQNRKLELSIMVVSROLVAVK----- 511
Db 465 PGGSHQLPEPTEELIECGKPLTVKRYTTEGEEKQFYQIMARNFTVDTGSEFEHFLIS 524
Qy 512 -----ONSTMSLTPEN--SW-----TPKACVIV 533
Db 525 EDKSGITDETYPIDVTALSLNPNBEPKENVVYIPPHIGETSLTIPSENNPASKILV 584
Qy 534 YVIEDDELIISDVKIPVOLVFNKIKLYMSKYKAPSEKVSLSISVTPDPSIVGIAYD 593
Db 585 FYVREDGETVADSTKITVKKLRNKGKFKGEKVLPGASSTLQTLAS-PYSICGIAVD 643
Qy 594 KSVNLMANASNDITMENNVNHEL-----ELNYTGY 622
Db 644 KSVHILSSDNRLTTEEVFNKLGHDYWPVKQATSDYKCEYKFKQTEGHEHGSFSGFT 703
Qy 623 LGMFNSFAVFOEGGLWLTLDANL-----TDYIDGVYDNAMEAEERFMEENEGHIV-- 673
Db 704 STNYIDSLITAFDEAGLVISDMELTRPCKSGFEDGRCQYDVAFAAPAARIGGG 763
Qy 674 -DIHDSIG-----SSP-HVRKHPETWIDTMNGYRIYQEFVYVPSDITSVATGF 725
Db 764 GEAGFGGQIRKKTNKPVEIRTYPPEITWELQNIAGTGLSLKRDIPHTITWVGSAT 823
Qy 726 VISEDGLGTLTPVELAQFQPFPLNLPYSVIRGEFALEITFNLYLKATKYKAVIIE 785
Db 824 CISEETGLGVSEA-ATVKGFOFFVFSFLPVSIVIGEKKPIIVTYFNYSLECLPIKLSIE 882
Qy 786 KSDKFDILMTSSEINATGHOQLVLPSEDAVVLPIRPTHLGEIPITVTALS----- 838
Db 883 QSDKF-----EMQNDINSYTSYCVCGKSGSDTRMMIRSLSGQVNLVYGLSLRNEAIC 935
Qy 839 -----PTASDAVTQMLVKAEGIEKSYSGSILLDLTDNRLOSTLKTLSFSPFPNTVYG 891

```

```

Db 936 GNQYVITYTAAADATROLVLVEBPGRPKEDMTSTACPEQDQNGKT-ATSULLPEDEVED 994
QY 892 SERVQITAI GDLVGPSINGLASLI RMPYGCGBOMNINFAPIYIILDYLTKKKQJTLNLIKE 951
Db 995 SARGVYSITGDLMPAIKNLMDLVRPLPTGCGEOMNWKVPVPIFVLDYLTATGSIITDSIKE 1054
QY 952 KALSFPMRGQYORELLIYQREDSFSAFGNYPDSGSTWLSAFLRCGLEADPIYIDONVHL 1011
Db 1055 KALNNMRGVAAROQNYRHPDGSYSAFGNRBDQGNLFLAFYRSFQAERPIILINKKYN 1114
QY 1012 RTYTMKJHOKSNBPFMDPGRVHSELQG---NKSPVTLTAIYVLSL---GYRKQPN 1065
Db 1115 ETENWILNRQSNQCFKRTIGLNSALKGGISSNDEPAPLTAIVLISLEAGTKN--- 1170
QY 1066 IDVOESIHFLESEFSRGIS-----DNYTLATITVALISGVSPPKAEALNMTLWRAEQ 1117
Db 1171 -----ETVIDQGISCEALSNPSTYSLALFAVATSLAGHPSADVLAKEBERAIT 1220
QY 1118 EGMQPFWVSSESKLSDSWQPRSLDI EVAAYV---LISHFLOFQTSSEGIPIKRWLSRQNS 1174
Db 1221 ECGGTFWKSPSS--GRYYWGNISIGVEIAGYAVLTLQHCGASNLAKVPIPIRLAKQONY 1278
QY 1175 LGFSASODTVALKALSEPAALNTERKTINQVYTWGSSPSPLAVUQ----- 1222
Db 1279 RGGFYSTODVTIALIQAMSKFATIIYKOBLEDEVGESSGFEKKIMLTQNSILMQTRFLO 1338
QY 1223 --PAVAINISANGFGPAICOLNVVNVYVKASGSSRRRSRIONGEAFDLVAVEN---KODLN 1278
Db 1339 TVBSPVDEPAEAGSGCGLVQTSIRYNVTP-----PPKGHLEVTYVRGILRDCIN 1389
QY 1279 HVDLNVCTSFSGPGR-SCMALMEVNLISGFVWPSBAIS-----LSFTVKVYVDHGKILN 1332
Db 1390 -AHIAATCVKYGDKGVSNMAVILEMKGVSGWIPDEBSIKNIVDREBLNLRREYVDONQNL 1448
QY 1333 YLDSVNEIQCVCNIPAVNPKYSNTQDASVSLVDYIERRQAVNSYNSEVLTSSCD 1388
Db 1449 YFSELTIDNLCEFNFWLEQDIEVOETKPTITLXYDYIELEGQVWVSYSD---ENCE 1501

```

```

RESULT 4
JC5143
alpha-macroglobulin precursor - guinea pig
C/Species: Cavia porcellus (guinea pig)
C/Date: 31-Jan-1997 #sequence_revision 27-Feb-1997 #text_change 20-Jun-2000
C/Accession: JC5143
R/Kwasaki, H.; Suzuki, Y.; Sinohara, H.
U. Biochem. J. 120, 1167-1175, 1996
A>Title: Cloning and sequencing of cDNAs encoding plasma alpha-macroglobulin and murineog
A/Reference number: JC5143; MUID:97164019; PMID:9010766
A/Contents: liver
A/Accession: JC5143
A>Status: preliminary
A/Molecule type: mRNA
A/Residues: 1-1476 <IWA>
A/Cross-references: DDBJ:D84338; NID:g1805591; PIDN:BAA12316.1; PID:g1304084
C/Comment: This protein belongs to alpha 2-macroglobulin family.
C/Superfamily: alpha-2-macroglobulin
F:1-23/Domain: signal sequence #status predicted <SIG>

Query Match      20.3%; Score 1491; DB 2; Length 1476;
Best Local Similarity 28.8%; Pred. No. 2.8e-76;
Matches 451; Conservative 292; Mismatches 524; Indels 300; Gaps 60;

Oy      4 PRLTAALLCCVCTALAVAPRRFLVTARPGIRRGANVTGVLLHHCBCQATYKAEI 62
    ||| | | | | | | | : : : : : : : : : : : : : : : :
Db      11 PVLL-----LTLILADAGASISGRPOVMVLVPISLHSGTPEKIQL-LTGLNETVTYKASLD 65

Oy      63 -LKTASNLTVASLEAGVEFEKGSFKTLTLPSLPUNLSADEIYEIVLRGRTODEILFSNSTR 121
    : : : : : : : : : : : : : : : : : : : : : :
Db      66 TIRENGSLPMMNVVAEKDLFOCVAF---TVPOSPIYEAMFLTVEVEGFTGH--FRSKKT 119

Oy     122 LSFETKRISIVEIQTDKALYKKQGEVKRFIVTLFSDPFKYKTSLN-----ILIKDPKSNI 176
    : : : : : : : : : : : : : : : : : : : : : :

```

Db	120	VLVSKXSLVFPQIDKRIYYRPGQTVKARVYSLDENRPP----	LNELPRLPIFDQPKGNRV	175		
Qy	177	QOMLSQSDJGVLSKTFQLSSHPILGWSIQV-QVNDQTYQSFQVSEYVLPRFEVTLQOT	235			
Db	176	MQOMNKLBERGLTQLSPLSEPLLSGYSVYVHKESGCRMHNSFTYBEFVLPRFEVQVM	235			
Qy	236	PLYSNANSKHLNGITTKATTKYQKPVKQDVTLLP-----	PLSTWGGKKKNTTKTF--	KING	288	
Db	236	PKKTLTLEQEFVSVCCRYTYGKVPNGNITMSICRNYNNSACLSEB----	SRAFCCKKYNQ	292		
Qy	289	SAN-----	FSFNDEMK-----	NMDSNLSLEYLDSFGPAELITTVT	328	
Db	293	QJNSQGFQIQVKTNDOQLRKREYENKLREAKIREBGT----	VOJLTGCFSEBITATTIT	348		
Qy	329	ESVYGISRNWSTNVFPKQHDYIIIEFPDYTVTLKPSLNFATKVTATRADGNQLTLEBRNN	368			
Db	349	K-----	LSFVKVDSYVARGVPFGQVRL--	VDGNVMPHPKM--	383	
Qy	389	VVITVTOQNTETWYSGNSGNQKMEAVQ-KINTYVPOSGTFKIEPFLIEDS----	ELQJ	443		
Db	384	ITITASEANY-----	HSNATTDENGLOFOSINTTMMIGSLINTQVKKHOSJNCYQWML	438		
Qy	444	KAYELGSKSSNAVHSLKPSKTYLOKT-----	RDNINVGSPFELVSGNKRJKE	495		
Db	439	EANEGASHTAAVPSL-----	SRSFVHLIEPOLGKLPCHQOTQFPA----	HYILNG-QELKE	489	
Qy	496	L--SYMVVSRQOLAVG-----	KONST-----	MFSLTPRENSMTPKACVITYYIEDDEIIL	543	
Db	490	LVFYVYIMAGGIVQSGTYVLSVQNGTKHFSVSYPVESDGLAPVARVLLYALLPSEIL	549			
Qy	544	SDVLKIPQVLQFNKKIKLYWSKVABESEKVSLSISTOPDSYGVIVADKSNVL-----	558			
Db	550	ADSAKYNVENCLDNKKNLSFSEGSPLASKTHLRVTAS--	POSICALPAVDQSVLLRKPEA	608		
Qy	599	-MANSN-----	DI-----	MENVHLELNTGTYIGMF-----	NMSFA	631
Db	609	VLSASVYALLPVKDLTGFPGLLQOEENQECVSLNLT--	YDGLIYSPREPNINEKQMYG	667		
Qy	632	VPOEGCLMVLLDANLTADYIDGVYDVAENAEYARFMEENEGHIVDIHPSLSSPH-----	685			
Db	668	FLXOMGLKVTNTKTIQKPOLCAHVQKEVTPMAVSSESS-----	SFSGPRPRVAVG	720		
Qy	686	-----	VAKHPETWIM--	LDTNMGRIYQEFVYVPSITSVAVTGFIYSED	730	
Db	721	IATYSEPPKETVATYSEPTWIMWLKATIDSSG--	VAEVEVTPDITTEWKAGAFCLISND	777		
Qy	731	LGIGITTPVELQAFQPPFIPLNPSYVINGEFALEITITFNKAKDTEVKVILIEKSKF	790			
Db	778	TGLIGISPL-ASLRAFQPPFELTTPSYVINGEAFLLAATVANYLAPDCIRISVHLEASP	836			
Qy	791	DIIMTSSINATGHOCTLAVSEBGAVLVPFIRPTHGELPIYVTA-----	836			
Db	837	LAERKAE-----	QESYCVQGNERQVSWVTPKSLGNVNPVYSABELESSELGNEKT	890		
Qy	837	LSPT--ASDAVTONILYKABGIEKSYQSILDLTDNRLQSTLKTLSFSFPNTVTGSEK	894			
Db	891	VVPFYGKKDITIKPLAVEPEGIEKEETWTSILIRVSDTTS--	EKLHELPSWVIOQSAR	947		
Qy	895	VQITAIQDVGPSINGLASLIMPMYGGGGEOMINFEANITYILDYLTQKKQLOTNLEKAL	954			
Db	948	ATVSIILGIIIGASAMQNTQNLQWMYGGGGEOMNYLFAENIVLYDLNTQOLTQLTPIKSKAI	1007			
Qy	955	SPMGQYQORELLYQREDGSPSAPG-AN--	DPSGSTWLSAFVURCLFLEADPYIIDQNVLHR	1012		
Db	1008	SYLSTGTOQRLNYNHRDQSYSTFEENRGGCGQNTWLAFLUKTFPSQARKIIFIDEAHITQ	1067			
Qy	1013	TYTMLKHQKNSGEFMDPGRVHSEELQGNKNSPVTLTAVYVTSLLGYRKYQPNIDVOESI	1072			
Db	1068	ALSWLSQKQKONGCFWSSGSLNNAIKGVYDELSISAVITIALLEMSLPDTHRVYNNAL	1127			
Qy	1073	HFLES-----	EFSGRISDNVTALLITYALSSVGS--	PKAKEALMMLTWRAEQSGMOPVY	1125	
Db	1128	FCLESAMKSAKEGTHG--	SHVYTKLLVYAPFALGNQKREIILKSLDEGVEKENSJLHWA	1186		



```

OY 1126 SSES-KLSDSW-----QPSLIDIEVAAYALLSHFLQFQISEIP-----IMRLSR 1170
Db 1187 RPQKREYSEGFKEFKSQAPSAEVMSTYVLLAYL-----YAPAPFPEDITSATDIDVNWYTK 1242
OY 1171 QNRSYSGASTODTTVALKALSEFAA--LMMTERTIQVTVTGSPSP-----LAV 1220
Db 1243 QONSHGCVSSYQDPTVALMALHSLSKYAAAFETFEKA--AQVTIKSGTFTNEVWNNRLL 1301
OY 1221 VQPMNV-----NISANGFPAICQLNVYVNV-----KASGSSRRRR 1256
Db 1302 LQOVSFLPVSQSYITITVYEGNVVYLQTSILKYNVSEKGTFFPALAEATVPQACGPKAHT 1361
OY 1257 SIQNGEAPFDLDVAVAKENDDLNHDVLDNYCTSPFGSGRSGMALLMEYNLLSGE--NVPS-EA 1313
Db 1362 SFQ-----ISLNVSYGSR--VSNMAYIVDKVSGTLPKPFYKN 1400
OY 1314 ISLSEYVKEVYEDHGKLMVLDVSVNETQFCVNIPAVRFKYSNTQDASVSIVDYEPBRQ 1373
Db 1401 LEKSHHSIRTEVSNHHVLIYIDKVSNOQLSLSPFVQDIEVRDLKPAITIKYDYETNEF 1460
OY 1374 AVRSYNS 1380
Db 1461 AIAEYHA 1467

RESULT 5
A23952
alpha-1 proteinase inhibitor III precursor - rat
C:Species: Rattus norvegicus (Norway rat)
C:Date: 19-Nov-1988 #sequence revision 19-Nov-1988 #text_change 16-Jul-1999
C:Accession: A23952; A27199; A34278
R:Brackish, T.A.; Northmann, W.; Hudson, G.O.; Shields, B.R.; Gehring, M.R.; Fey, G.H.
J. Biol. Chem. 263, 3999-4012, 1988
A:Title: Sequence and acute phase regulation of rat alpha-1-inhibitor III messenger RNA
A:Reference number: A27709; MUID:88153707; PMID:2831216
A:Accession: A23952
A:Molecule type: mRNA
A:Residues: 1-1477 <BRA>
A:Cross-references: GB:J03553; NID:g202571; PIDN:AAA40628.1; PID:g202572
R:Schweitzer, M.; Takabayashi, K.; Geisler, T.; Laux, T.; Biermann, G.; Buhler, J.M.; Gauts
Eur. J. Biochem. 164, 375-381, 1987
A:Title: Identification and sequencing of cDNA clones for the rodent negative acute-phase
A:Reference number: A27199; MUID:87190405; PMID:2436907
A:Accession: A27199
A:Molecule type: mRNA
A:Residues: 1253-1477 <SCH>
A:Cross-references: GB:M8297; NID:g204930; PIDN:AAA63493.1; PID:g204931
A:Experimental source: Liver, clone p13cDNA6
R:Englind, J.U.; Salvesen, G.; Thøgersen, I.B.; Pizzo, S.V.
J. Biol. Chem. 264, 11428-11435, 1989
A:Title: Proteinase binding and inhibition by the monomeric alpha-macroglobulin rat alpha
A:Reference number: A34278; MUID:89291900; PMID:2472396
A:Accession: A34278
A:Molecule type: Protein
A:Residues: 673-676, 'T', 678-723 <ENG>
C:Superfamily: alpha-2-macroglobulin
F:1-24/Domain: signal sequence #status predicted <SIG>
F:25-1477/Product: alpha-1-inhibitor III #status predicted <MAT>

Query Match 19.8%; Score 1455; DB 2; Length 1477;
Best Local Similarity 29.4%; Pred. No.3.1e-74;
Matches 457; Conservative 237; Mismatches 578; Indels 284; Gaps 56;

OY 13 LCVCMAALNAPGPRFLVTAPOGIIIRPGN---VTIGVLLHCHPQSVTVKAEILLKTASN 68
Db 9 LCLPSALALPFPASLL-----NGNSKRYMLVPSQLTETPEKICMLYHINETVT 59
OY 69 LTVSLVLEAG-----VFEKGSFK--TLTLPSPPLNSADEIYELRVYGTQDEILFSN 118
Db 60 VTASISISGRGIRKLFDLIVVDKDLFHCYSPFITIRPLPSEEEESLDINIEGAKH---PSE 116
OY 119 STRLSFETKRIISVFIQTDKALYKPKQEVKFRIVTLFSDFKPKYKTSNLIL-----IKDPKS 173

```

Dd	117	NRVLVAKKESVVAVQTDKPMYTKGQVKKRVVSMKDLHP----	INLPLPLAIEDPKM	172
Qy	174	NLIQWMLSQOSDLGVISKTQFQSLSPHLLGDMSI-----	QOVNDQTYQOSFOVSEVYLPKF	229
Dd	173	NRIMQODVKTENGLKQLSFSLSAEPILQGPYKILVILQSGVKEB--	HSFTVMEFVLPRF	229
Qy	230	EYVLQTLPYGMSKHLNGITITACTYTGKPYKGDVYTLFLPLSPFKKKNIITFKINGS	289	
Dd	230	GVDKVPNAISVYDEIINVTACATYTGKPEVPGVKISL-----	CHGN	272
Qy	290	ANFSFNDBEKMNVDLS--SNGLS-----	EYDLSPPGPVEILITVTESVGT	333
Dd	273	PTFSEETKSGCKEDSKLDNNGSGTQVNIITERQKENTYKMQO--	AFHVAITVTEGCTG	330
Qy	334	ISRNVTNVFQKHDIIEFFDYTYLKPSLNFATVATVKTADGNQUTLLERRNNVITY	393	
Dd	331	SEFGSGRIEVERTRNKFLLKADSHRPHGIPFVVKRVLVDIKGDPINBQ----	VLKA	366
Qy	394	TQRVITYEWSGNSNGNQMEAVQKINTVYQSGTFKIEFFILEDSSLOLKAYFLSGKS	453	
Dd	387	RDAGYTNATTTDQHLAK-----	FSIDTNGISDYS--LNTKVH--KEES	427
Qy	454	MAVVS-----	LFKSPSKTYIQLKTR-----DENIKGSPPEL--VYSGN	490
Dd	428	SCIHSSCLAEHNAHHTAAVAVSLSSYIYLTDEAGVLEPCNQIHTYQAFILKGVLGV	487	
Qy	491	KRLKELSYWVSRGQLAVGK-----	QNSTMFLTPBNSMTPRACVIVYYIED	538
Dd	488	LQQIVFHYLVAQSSILQTDGNHTQVBPESGQVGNFALPIVEFSWVPVAKMLIYTLIP	547	
Qy	539	DGEIISVULKIPVOLVPRKIKILYMSKVKAPEPSKVSLSISTVQPSDITGIVAKRVNL	598	
Dd	548	DGEVIAVSVKFQVKEKCLRNKVHLSFSQSSQSLPASHQTHRKVTAS--	POSLCGELRAVDOSVLL	606
Qy	599	-----	MNASNDITMBENVVHELELYTNGYV--LGMFNNSFAVEFOECQAMV	640
Dd	607	QKPEALISPSLIVDLPQMODSNFLASNDPFEBDEDYCLMTQPLAREKDVRYRARETILMA	666	
Qy	641	LTDANL--TKQIDGVYNAEYA-----	ERFMEENEGHIVDIDHPSLGSSP-----	684
Dd	667	FTNLKILPFCYCNTRYDMVPLAVPAVALDSDSTRGMYESL--	PVAVKS--PLPOEPPRPKD	724
Qy	685	-----	HYRKHEPEWTIMDLDTMGGRIRYOEPFVTVPDSTISMVANGVISEDGLGT	736
Dd	725	PKQDVIETIRKNEPEWTIMDLVTVSSGVTLEMTVPDITTEKACALCLSDNTGLGLS	784	
Qy	737	TTPELOAFOPFFIFLNPYSVIRGEEFALETITFNVLKQATEKVIIEKSDKFDLIMTS	796	
Dd	785	SV-ASPAQFOPFELMFPYSVIRGEAFILKATVANTLPLSLMAVULEBSPPF-----	837	
Qy	797	SEINATGHQOTLVPSDDGATVLPPIRPHLGEIPITVIALS-----	PTAS-----	842
Dd	838	TAVENVNODSYCLGANGRHITSGLVLPKSLGVNVEFSAEAAQSPQCSSEVATVPETG	897	
Qy	843	--DAVTQMILVKEGIEKYSQSILDLTDNRLOSTIKTSLSPFPPTVYGSRRVOITAI	900	
Dd	898	RKDTVAVYLVLEPBGIKKHTFSSLLCASPAELS--	ETLSLLPLPVVXDSAAHPVM	954
Qy	901	GDVLGPSINGLASLIRMPYCGGEONMINFANITYIIDLTKKKQOTDNLKEXALSFWROG	960	
Dd	955	GDILSSAIKMTQNLIQMPYGGGEGQNMVLPAPNITYVLKYNLETQOLTEKISKALGYLRAG	1014	
Qy	961	YORELLAYOREDGSASARGND--	PSGSTMASATPLRCFLPADRIDQNVLAHRTYTWLK	1018
Dd	1015	YORELANTKHODGSASAGDHNGOQGNWTMLTAVLKSFAQARAFIFIDESHITDAFTWLS	1074	
Qy	1019	GHQSGNEFMDPGVHISELOGGKNSPVTLTAYIVLSLIGYRKQOPNID--	VOESIHFLF	1076
Dd	1075	KQQDSCGCFSSSGSLNNAKKGAVDEITISATITYALL--	ESSLPDTPVYSKALSCLF	1132
Qy	1077	SEFSR-----	GISDNYTLALITYALSVSQ--PRAKELANMLTVRABQGGMOFW-----	1124

Db 1133 SSWENIEOGNGSFVYTKALMAVAPALAGNOEKNEILKSLDKAEIKEDNSIHMERPOKP 1192  
 QY 1125 VSSSEKLSDSWQPSRLDIEVAAYALLSHFQFOTSEGP-----IWRMLSPRONS 1174  
 Db 1193 TKSGCYLYTP-QASSAEVEMSAIVYLARL-----TAQPAPESEDIALSMGTIKMLTKOONS 1247  
 QY 1175 LGGFASITQDTTVALKALSER-AALMNTERTNIQVTVTPSSPSPLAVVQ----- 1222  
 Db 1248 YGGRSSITQDTVALDALSKYGAAIFSKSQKTPSTVYSSGSGFQKQFQVDSNRLLLOQVS 1307  
 QY 1223 ----PMAVNISANQFGAICQLNVTNYVYKASGSSRRRSIONQAFDLDAVK---ENK 1274  
 Db 1308 LPYIPGYVTVYVSESGCYAQTTRRYNPLE-----KQAPAFALKYQTVPLTCNMP 1358  
 QY 1275 DDLNHLVNLVCTSPSGP-GRSGMALMEVNLISGFVMSSEALSSETVYKV-----E 1324  
 Db 1359 KGQNSFQISLEISTYMGSRPASNMYIADKMLSGF-----LPLKPTAKKERLCHVSRTE 1412  
 QY 1325 YDHGKMLYLDPSVNEITPCVNIIPAVRNFKKYSNTODASVSIYDYVYPRROAVRSYNS 1380  
 Db 1413 VTTNNVLLYLDQVNTQNTLSFSEFIILQODIPVKNLQPALYKVYDYETDEVAFAEYSS 1468

## RESULT 6

MAHU  
 alpha-2-macroglobulin precursor [validated] - human

N/Alternate names: alpha-2M  
 C/Species: Homo sapiens (man)  
 C/Date: 05-Apr-1993 #sequence\_revision 30-Jun-1987 #text\_change 08-Dec-2000  
 C/Accession: A94033; S09107; JN0262; A92486; S6634; A01256  
 R/Kan, C.C.; Solomon, E.; Belt, K.T.; Chalm, A.C.; Hiron, L.R.; Fey, G.  
 Proc. Natl. Acad. Sci. U.S.A. 82, 2282-2286, 1985  
 A/Title: Nucleotide sequence of cDNA encoding human alpha-2-macroglobulin and assignment  
 A/Reference number: A94033; MUID:85190481; PMID:2581245  
 A/Accession: A94033  
 A/Molecule type: mRNA  
 A/Residues: 1-1474 <MAN>  
 A/Cross-references: GB:M1133; NID:G177869; PIDN:AAAS1551.1; PID:G177870  
 A/Note: hydrolysis of the thiolester bond during amino acid sequencing of the mature pro  
 lymorphism  
 R/Bell, G.T.; Rall, L.B.; Sanchez-Pescador, R.; Merryweather, J.P.; Scott, J.; Bddy, R.L  
 Somat. Cell Mol. Genet. 11, 285-289, 1985  
 A/Title: Human alpha-2-macroglobulin gene is located on chromosome 12.  
 A/Reference number: I39375; MUID:85219061; PMID:2408344  
 A/Accession: I39375  
 A/Status: preliminary; translated from GB/EMBL/DBJ  
 A/Molecule type: mRNA  
 A/Residues: 832-999; 'I', 1001-1147, 'D', 1149-1194, 'D', 1196-1474 <BEL>  
 A/Cross-references: GB:M4501; NID:G177871; PIDN:AAAS1552.1; PID:G177872  
 R/Marynen, P.; Devriendt, K.; van den Berghe, H.; Cassiman, J.J.  
 FEBS Lett. 262, 349-352, 1990  
 A/Title: A genetic polymorphism in a functional domain of human pregnancy zone protein:  
 n.  
 A/Reference number: S09106; MUID:90242963; PMID:1692292  
 A/Accession: S09107  
 A/Status: not compared with conceptual translation  
 A/Molecule type: DNA  
 A/Residues: 672-747 <MAR>  
 R/Matthijs, G.; Devriendt, K.; Cassiman, J.J.; Van Den Berghe, H.; Marynen, P.  
 Biochem. Biophys. Res. Commun. 184, 596-603, 1992  
 A/Title: Structure of the human alpha-2-macroglobulin gene and its promoter.  
 A/Reference number: JN0262; MUID:92246939; PMID:1374237  
 A/Accession: JN0262  
 A/Status: not compared with conceptual translation  
 A/Molecule type: DNA  
 A/Residues: 1-29 <MAT>  
 R/Sottrup-Jensen, L.; Stepanik, T.M.; Kristensen, T.; Wierzbicki, D.M.; Jones, C.M.; Lott  
 J. Biol. Chem. 259, 8318-8327, 1984  
 A/Title: Primary structure of human alpha-2-macroglobulin. V. The complete structure.  
 A/Reference number: A92486; MUID:84239807; PMID:6203908  
 A/Accession: A92486  
 A/Molecule type: protein  
 A/Residues: 24-562, 'E', 564-974, 'Z', 976-999, 'I', 1001-1474 <SOT>

R/Sottrup-Jensen, L.; Stepanik, T.M.; Kristensen, T.; Wierzbicki, D.M.; Jones, C.M.; Lott  
 J. Biol. Chem. 260, 6500, 1985  
 A/Reference number: A92529  
 A/Contents: annotation; extractum  
 R/Virca, G.D.; Salveen, G.S.; Travis, J.  
 Hope-Seyler's Z. Physiol. Chem. 364, 1297-1302, 1983  
 A/Title: Human neutrophil elastase and cathepsin G cleavage sites in the bait region of  
 A/Reference number: A91713; MUID:84030513; PMID:6195065  
 A/Contents: annotation; inhibitory site  
 R/Sottrup-Jensen, L.; Lombard, P.B.; Stepanik, T.M.; Petersen, T.E.; Magnusson, S.; Jor  
 FEBS Lett. 127, 167-173, 1981  
 A/Title: Primary structure of the 'bait' region for proteinases in alpha-2-macroglobulin  
 A/Reference number: A91290; MUID:81212827; PMID:6165619  
 A/Contents: annotation; inhibitory site  
 R/Hall, P.K.; Nelles, L.P.; Travis, J.; Roberts, R.C.  
 Biochem. Biophys. Res. Commun. 100, 8-16, 1981  
 A/Title: Proteolytic cleavage sites on alpha-2-macroglobulin resulting in proteinase binc  
 A/Reference number: A90099; MUID:81255805; PMID:6167263  
 A/Contents: annotation; inhibitory site  
 R/Mortensen, S.B.; Sottrup-Jensen, L.; Hansen, H.F.; Petersen, T.E.; Magnusson, S.  
 FEBS Lett. 135, 295-300, 1981  
 A/Title: Primary and secondary cleavage sites in the bait region of alpha-2-macroglobulin  
 A/Reference number: A91299; MUID:82056610; PMID:6172288  
 A/Contents: annotation; inhibitory site  
 R/Dolmer, K.; Jenner, L.B.; Jacobsen, L.; Andersen, G.R.; Koch, T.J.; Thirup, S.; Sottrup  
 FEBS Lett. 372, 93-95, 1995  
 A/Title: Crystallisation and preliminary X-ray analysis of the receptor-binding domain of  
 A/Reference number: S6634; MUID:9602553; PMID:7556651  
 A/Accession: S6634  
 A/Molecule type: protein  
 A/Residues: 1337-1343 <DOL>  
 C/Comment: This inhibitor is able to inhibit all four classes (EC 3.4.21-3.4.24) of endo  
 g tertiary conformation, entrapping the proteinase. This results in hydrolysis of its thi  
 e and appears to unmask a receptor-specific binding site.  
 C/Comment: The entrapped enzyme remains active against low molecular weight substrates (e  
 r complex results in its rapid clearance from the circulation by receptor-mediated endoc  
 C/Comment: The wide specificity of this inhibitor is attributed to the primary sequence c  
 ities of a variety of plasma proteinases, form primary (residues 704-809) and secondary  
 C/Generics:  
 A/Gene: GDB:A2M  
 A/Cross-references: GDB:119639; OMIM:103950  
 A/Map position: 12p13.3-12p12.3  
 C/Complex: homotetramer; dimer of disulfide linked dimers  
 C/Superfamily: alpha-2-macroglobulin  
 C/Keywords: glycoprotein; plasma; proteinase inhibitor; thiolester bond  
 F.1-23/Domain: signal sequence #status predicted <SIG>  
 F.24-1474/Product: alpha-2-macroglobulin #status experimental <MA2>  
 F.693-694,704-709,719-723,730-735/Region: inhibitory #status predicted  
 F.698-742/Region: bait region  
 F.48-86,251-299,269-287,278-431,595-771,642-689,821-849,847-883,921-1321,1079-1127,1352-  
 F.55,70,247,396,410,869,991,1424/Binding site: carbohydrate (Aen) (covalent) #status exp  
 F.470,563/Disulfide bonds: interchain #status predicted  
 F.972-975/Cross-link: thiolester (Cys-Gln) #status experimental

Query Match 19.6%; Score 1440; DB 1; Length 1474;

Best local similarity 29.1%; Pred. No. 2, 2e-73;  
 Matches 451; Conservative 266; Mismatches 551; Indels 284; Gaps 59;

QY 12 LUCVTAALAVADPPEFLVTAPGIRPGANVTIGVELHCESQVTKAEELKTASNLTV 71  
 Db 15 LVALFLPTDASVSGKPOQMYLVPSLHRT-ETTEKGVCLLSLNTVNTVSSLSBVRGNRL 73  
 QY 72 -SVLEAG-VFEKSGKRTILTPSLPNSADE---IYELVVTGTODEILFNSNTRLSFET 126  
 Db 74 FTDLAEADVLCVAF-----AVPKSSNSEEVMLPTVOVKGFQD---FKKRTYVAVKN 124  
 QY 127 KRISVLEIDTKALKYKKEVFRIVTLFSDPKFKYKTSNLT-1KDKSNLIQOMLSQSQSD 185  
 Db 125 EBSLVVQIDKSLYKKGQYTKFRVSMDSNFHNLNLPVLYIQDPKSGRIQAQMSQGLE 184  
 QY 186 LGVISTPQLSSHPILIGDSIQVQND--QTYQSQVSESYVLPKREVTLCPLYSQMS 243  
 Db 185 GGIKQSPFLSSPFGSGYKVVQKKSGRTEH-PTVEBEFVLPKFEVQTVVPKIITILE 243

QY	244	KHLNGITLAKXNYGKRVKGDYTLTFLPL-----	SPFKG	276
Db	244	EENNVSVGCLYYTGKVPFGHVTIVSICRKYSADSDCHGEDSOAFCEKFSGOLNSHCFCFQO	303	
QY	277	KKNITKTFKINSANFSPFNDEMKKVMDSNGLSEYLDSPSGPEILLT-----	TVTSV	331
Db	304	VR--ITKPOLKR-----KEYEMK-----	LHBAQIQEBSGYVELTGQSSSITTYI	347
QY	332	TGISRNVSINVFKOHDIYIEFPDYTYLKLPSLNFATATKVTYRADGNQTLJEBRRNNVI	391	
Db	348	TKLS-FVKVDSHFROQ--IPFFG-----	QVRLVDGKGVP-IP--	384
QY	392	TYTQRYRTYYSGNSNGNQKMAVO-KINYT--VQSGTFKEIRP-----	ILEDS	438
Db	385	FI-RGNEANY--SNATDEHGLVQFSINTTVMGSLTVRVNNYKDRSPCYQWVSREH	441	
QY	439	SELOKAYFLGSKSSVAHSLFKPSKTYIOKTKDENIKVGS----	PFELVSGNK--R	492
Db	442	EEAHHTAY-----LVSPSKSVHLEPMSHEPCCHTQTOAHYILNGTLLG	489	
QY	493	LKELS--YVWVRGOLVAVG-----KONST--MFSLTPENSWTPRACYIVYIEDG	540	
Db	490	LKLTSLYYLIAMKGGIVRTGTHGLVKQEDMKGHRISISIPKSDIAPARLLIYALVPTG	549	
QY	541	EIISDVLKIPVOLVFNKIKIYWSKVKAPSEKSLKISVTPQDSIVGIVAVDSVNLNM	600	
Db	550	DVIGSAKYDVENCANLKNVDLSFSSQSLPASHAHLRVTA--PQSVICALRAVDOSVILMK	608	
QY	601	ASNDITMEVNHLEL-ELVNTGY-----YLGMPNS-----	FA	631
Db	609	PAELSSASSVYVLLPEKDLTGPPGLNDODEDCINRHNVYINGITTYTPVASTNEKOMYS	668	
QY	632	VEQECGLMWLTJANLTOKYIDGVYDNAE-----YARFMEBEGHVIDHDPSL	680	
Db	669	FLEDGKLKFTMSKIRKPKCMCPQLOQYEHNGEGELRVGFYSBDVNGRHALVHVE--	725	
QY	681	GSSPH--VRKHPETWIMLDTNMGYRIYQEFVTPDPSITSWATGVFVISEDGLGLTT	737	
Db	726	---PHTEYRKXFPETWIMDLVWVNSAGVAVGVTPPTITBWKAGARCLSEBDGLGIS	782	
QY	738	TVEVLQAPQPFITPLTPSVYRGEFPALETITNYLKDATBEVKYIIEKSDKFDILMTSS	797	
Db	783	T-ASLRAPQPFVEVLTMPYSVIRGEAFLLKALVNLVLPICIRVSVQLEASPAFLAVPEK	841	
QY	798	E-----INATGQOQTLVSEBDGATVLPDIRTHGEIPITYTALS-----	PTA	841
Db	842	EOAPHCICANGQ-----TVSAVTPKSIANNVFTVSAAALSSQELCTGEVDSV	890	
QY	842	-----SDAVTOMILVKAEGIEKYSOSILDLTDNRLOSTLKTISFSPFPYVTGSEKVO	896	
Db	891	PEGRKRDYIKPLVLEBPELEKETTFNLSLCSGGEVS---BELSLKIPRVVSESAAS	947	
QY	897	ITALIGDVLPISINGLASLIRMPYGGEBQNMINFANIIYLDYLTKKQLTNNLEKALISF	956	
Db	948	VEVLGDIIGSAMQNTONILQMYGCGEOMVFAFNIYVLDYANLTQOLTPEVKSKAIGY	1007	
QY	957	MKGQVORELLVQREGEFSAPG-ANDPS-GSTWLSAPFLRCLEADPYIIDQVNLHNTY	1014	
Db	1008	LMTGTYOROLNHYKHGDSYSTFERGRGQNGNTWLAFLKTPAQARAVIIFIDEAHITQAL	1067	
QY	1015	TWLKHQKSNCEFWMDPGRVHSELQGNKSPYTLATVYVTSLSGRKYOPMIDVOESIH	1074	
Db	1068	YLSQROKONGCFRSSGSLNNAIKGVDEVTLSAYITIALLEIPLVTHPRVNNALFC	1127	
QY	1075	LBS-----EFSRGISDNTYLLITYALSSVGS-PRAKEALNMLTWRAEQEGNQFW--	1124	
Db	1128	LESAMKTAOEGDHG--SHVYTKALLVAFALAGNQDKREVKLSINEEAVKDNQSVHWRP	1186	
QY	1125	VSESEKLSWSQPR--SLDIEVAAYALLSHFIQO--TSEIP---IMWLSQRNSL	1175	
Db	1187	QKRAVGVHFEYEQADSAVENTSVYLLA-YTTAPAPISDEUTSATNIVMTITQQDNAQ	1245	

QY		1176	GGPASTOOTTAALKALSEFALMTERTINIQVTWGPSSPSPLAVVQ-----	1222
Db		1246	GGSFTSTDTVALHNLKSKGATTF-KTGAQAQVITGSGTFSRFQVDNNRLLAQVS	1304
QY		1223	---PMANINISANGGFAPICQLNVVVYNKASGSSRRRSRSIONOEAFFDLDAVK--ENKD	1278
Db		1305	LPELGEYSMKVYGCCYYLTQTSLKVN-----LPKEFPFALQGLTLPGTCB	1355
QY		1276	DLN-HVDLVNCTSPSGPC--RSGMALMEVNLISGMVPSEAISL--SETYKKVEYDHG	1322
Db		1354	EPKATTSFOILSLSYSTSRASRNMAIIVDKVWSGFIPLKPTVMLEKSNHVSRTREVSN	1411
QY		1329	KLNLYLDSVNETGPCVINPAVRNFVNSTODASVISUYEPRRAVNSYS	1380
Db		1414	HVLIIYLDKVSNGTUSLFPTTVLDQVPRDLKPALIKVDYETDEFALAEYNA	1465
<b>RESULT 7</b>				
		T43166		
		alpha-2-macroglobulin - Japanese lamprey		
		C;Species: Lampetra japonica (Japanese lamprey)		
		C;Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 20-Jun-2000		
		R;Nonaka, M.		
		submitted to the EMBL Data Library, November 1992		
		A;Accession number: Z22323		
		A;Accession: T43166		
		A;Status: preliminary; translated from GB/EMBL/DBJ		
		A;Molecule type: mRNA		
		A;Residues: 1-1503 <NON>		
		A;Cross-references: EMBL:D13567; PIDD:BA02762.1		
		A;Experimental source: liver; clone Cl.3 and Cl1.10		
		C;Superfamily: alpha-2-macroglobulin		
	Query Match	19.4%; Score 1424; DB 2; Length 1503;		
	Better Local Similarity	28.3%; Pred. No. 1.8e-72;		
	Matches	442; Conservative 256; Mismatches 600; Indels 266; Gaps		
QY		10	AHLTCVCAALAAVAPGR---FLVTAAGIRPGGNVTIGVELHEHCPSQVAKAELKT	65
Db		7	AFULLLCMLAAVSADEBSGHVLPVPSLHALSSELCSLL-GVIGVTFPRATLHYK	65
QY		66	ASNLTVSYLEABGVFEKGSFKTLTPS-----LPLNSADEIELRVGTQDEI	114
Db		66	DSRDTSVAQHNTTVGEVDACVLLQPSGDLHGCFSTVP-DVGQTYANLVRAAGGL	124
QY		115	LESNSTRLSFETKRISVFIOFDKALKYRKQSVKRIYTLFSDFKR-YKTSINLIKDPKS	173
Db		125	NFTKTHAVVARKVDVVFQTDKPKYKKGQSVKFRVWLIDENFAVLTKYALITYIEDQR	184
QY		174	NLIQOMLSQSGDLAGISKTFOLSSHPIIDGMSIOY---QVNDQFYVSGFOYSEYVLPKE	230
Db		185	NRIAMKRNASGRAGIVQLELMIPSEPPLGTYNVNWVEBSGDVNASHTFYBEYLPTFE	244
QY		231	VTLQTPILCSMNSKRLNGTIYAKTYTGKPVXGVDTLFLPLSPWGKKXNIYKTFKINGSA	290
Db		245	VSIGQPSYANTLADKSVTLVKCGARTYTGKPVGAVVASVC-----IQQP	288
QY		251	NFSPPDEMKRMVD-----SSNGLSER-LULSPG-----PVELLTVTESVTGIS	335
Db		289	RFWMTEECILPVNCNEFFPMKVGKDGCAEQVNAKRSNASCHTNHKLKVAAALEEGTQM	348
QY		336	RNVSTNVFFKHQDYIEFFDDYTVLKPSLNFATATKYVTRAGNOQLTERNRNVVITYQ	395
Db		349	MKQAAEKPFEDDIRISFVDMPSWYRHSLPIYGVKAYKERPGOSPVP---HKLVSLIYQ	404
QY		396	RNYTEWMSGNSNQKWEAVQKINYVPOSGTFKIEPILDESSELQIKAYFLAG--KS	452
Db		405	GNAAPRGSHHTG----ADGTFETLI-DTGDFNRSDTIFLEANDPEFNSTAHPVTVQQ	458
QY		453	SMAVHSLEKFSKSYIYIQLKTDENIKGVS--PFELVYSGNKRLK-----	495
Db		459	GYSITISAFSSDSDELQDRVAHTLEGCSSVPLRLILV---LKENRSSAGDHAGVGH	514

```

QY 496 --LSYVVSRCOL-----VAVGKQNTMFSLPENSWTPKACVIVYIEDDEII 543
D 515 PVINVLVMSRGNIIHTESFTMDNVYLKGAKSHHFOQKSHFAPARLLAMVFN--ETV 573
QY 544 SDVAKIPQVLPKPKIKIYMSKVAEAPBEKSLISVYQPSIYGIYAVVKSVMVMAASN 603
D 574 ADTIALPVHKCFPKPKVSVGFEBEDELPGAMSTLIRAS--PSLALRVVDSVILLPREA 632
QY 604 DITMENNVHLELYNTGY-----LCMFNNSF-----AVFORGL 638
D 633 QLSAESITQRLQVODLTYYDYGIERMDCSPHYRRKRKRSYVMWDDFQNSVHRLFKMGGL 692
QY 639 WVLTDANLTQDYID-----GVYDAAEYAEFMEENEGHIVDHFSLGSS-- 683
D 693 VVLIDTTV--KRSVDCQHMHRHRLIAYSMETVAPSGARFSDRG-----NISPVGGGMF 746
QY 684 ----PHAKHPEYVIMLDTMWGRIRYOEFEVYTPDSTTSVAVAGFVISELGLITTP 739
D 747 KTVKETVREYEPETWIMDLVPSVSGLEQVAVKVPDSITTEKASAF--CSSPAGFGLSEVS 805
QY 740 VELAPOFFFIPLNPEYVIRGEPALEITTFNYLKATEVKYIIEKSDKDIIMTSEI 799
D 806 --SLKVTFFPEYEPVLPYSVAGETFPPLAISVANTLHSCLEVTLLDSEHFAV----- 857
QY 800 NATCHQOQ---TLVPSDEGATVLPPIRETHLGEIPITVVALS-----PT 840
D 858 --AAGSPRAASCVCPRKRG--LFSFLIEPLALGTNVVSRAAVAHSDPCGNEVVVVEEG 915
QY 841 ASDAVTOMILYKAGIEKYSQSLIDLITDNRLQSTLKTSSFSFPTVTGSESVQITAI 900
D 916 AVDTVSVSVIPEPPIPELAVYSL--LCPRKSPASERFVFLPLPAVVIYSARAAYTIA 973
QY 901 GDVIGPSINGLASLIRMYGCGEONMINEFNIYILDVLTCKKLTNLKKAISPMROG 960
D 974 GDINGSAQNDKLLTLPFGCGEONMVAFAFNITIOELQNSGQLTAVRKALNPLRVG 1033
QY 961 YORELTYOREDESSAFGNYPDPSGTMWLSAFVLCFLEADYIDIDONVHRTYTWLXGH 1020
D 1034 YORQLTYKRDHSHSARFGKSDDDONTWLTAVLVSFPASGHIAVSEDHITGPRSWLVEH 1093
QY 1021 QK--SNGEWMDPGRVYHSELQGNKSPVTLNVIYTSLSIGRKYQPNID--VQESI-- 1073
D 1094 QNASTGCFISVGRFPNNAMKGVSDVDVSLTAYVTRAL--ESNISGPVVERALCLR 1148
QY 1074 --FLESERSRGISUNYTLALITYALS--SVGSPKAKELNMLTWBADEGCGQFVSESK 1130
D 1149 PLVLES-----SNAHLIALASVAFPSLSADGATQALRDALQSRATVTKGSLTHQZANKK 1202
QY 1131 LSDSWQPR-----SLDIEVAAYALS-----HFQFQTSBEGIPIMRLSR 1170
D 1203 DDKQEBDEEGFRRRSYGTTSAAVENTYALALLSRALVPOGLAASATSSNIGVQWLK 1262
QY 1171 QNSLGGFASPTODTTVALKALSEPAL-----MNTERTNI--QVTVGP 1212
D 1263 QRNAVGGSSPTQDTVVGGLQALSAFAALIDGCGGSGSGIKAVVLLDAHHTLREVSIDST 1322
QY 1213 SS-----PSPLAVQPMAMVINSANGFPAIQOLANVYVYKASGSSRRRSIQNEAF 1264
D 1323 NALLIHQVLPFPVAPWVSSCTVEATGQCALQVSLKNEPPKSKPK-----F 1373
QY 1265 DLDAVAVK--NKQDLNHDVLDNCTSPSGP--GRSGMLMEVNLGSPMVPSEALSLSET-- 1319
D 1374 SLSTETNPANCSKAAKRGQJNVEVSHYGERGEMMALVEKMTSGISAYKSSLSKEQITF 1433
QY 1320 --VKKVEYDHGKLTNLYDSVNETQFCVNIIPAVENFKVSNTOASVSIVDYEBRQAVR 1376
D 1434 YDFYKATVEVDASRVVYLDKVDKXPIALKISVTQDIADVNDLQPATVAVDYDATEDATS 1493
QY 1377 SYNS 1380
D 1494 MYSA 1497

```

## RESULT 8

```

S15904
Alpha-1 protease inhibitor III, variant 1 precursor - rat
C:Species: Rattus norvegicus (Norway rat)
C>Date: 30-Jun-1992 #sequence revision 30-Jun-1992 #ext change 16-Jul-1999
A:Accession: S15904; S52235; S29737; B34278; B29952; C29952; B27199
R:Regier, R.; Stickinger, S.; Schweizer, M.
FEBS Lett. 282, 368-372, 1991
A>Title: Differential regulation of the two mRNA species of the rodent negative acute phase
A:Reference number: S15904; PMID:91243832; PMID:1709877
A:Accession: S15904
A:Molecule type: mRNA
A:Residues: 1-1487 <REG>
A:Cross-references: EMBL:X52984; NID:g55561; PIDN:CAA37176.1; PID:g55562
R:Stickinger, S.; Schweizer, M.
submitted to the EMBL data library, May 1992
A:Reference number: S52235
A:Accession: S52235
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-161 <SIC>
A:Cross-references: EMBL:X66454; NID:g609245; PIDN:CAA47070.1; PID:g609246
R:Thomsen, N.K.; Soltrop-Jensen, L.
Arch. Biochem. Biophys. 300, 327-334, 1993
A>Title: alpha-macroglobulin domain structure studied by specific limited proteolysis.
A:Reference number: S29737; PMID:93143334; PMID:7678727
A:Accession: S29737
A:Molecule type: protein
A:Residues: 25-32,276-280, 'XG',314-320,'X',322,'SL',717,'VAV',722-726,'Q',728,1205-1218
R:Englhd, J.C.; Salvesen, G.; Thogersen, I.B.; Pizzo, S.V.
J. Biol. Chem. 264, 11428-11435, 1989
A>Title: Proteinase binding and inhibition by the monomeric alpha-macroglobulin rat alpha
A:Reference number: A34278; PMID:69291900; PMID:2472396
A:Accession: B34278
A:Molecule type: protein
A:Residues: 684-715,'N',717-733 <ENG>
R:Brackak, T.A.; Northam, W.; Hudson, G.O.; Shleis, B.R.; Gehring, M.R.; Fey, G.H.
J. Biol. Chem. 263, 3999-4012, 1988
A>Title: Sequence and acute phase regulation of rat alpha-1-inhibitor III messenger RNA.
A:Reference number: A92709; PMID:68153707; PMID:2831216
A:Accession: B29952
A:Molecule type: mRNA
A:Residues: 1-3,'D',5-86,'V',98-136,'M',138-169,361-364,'M',366,'V',368-397,'H',399-727,'
A:Cross-references: GB:J03552; GB:M22359
A>Note: the authors translated the codon ATG for residue 174 as Ile
A:Accession: C29952
A:Molecule type: mRNA
A:Residues: 80-262,269-361 <BR2>
A:Cross-references: GB:M22360; NID:g202583; PIDN:AAA40633.1; PID:g554402
A:Accession: D29952
A:Molecule type: mRNA
A:Residues: 315-409,1064-1079,'LAASGAPDHC',1090,'TMP' <BR3>
A:Cross-references: GB:M22361; GB:M22362
R:Schweizer, M.; Takabayashi, K.; Geiger, T.; Laux, T.; Biermann, G.; Buhler, J.M.; Gautt
Eur. J. Biochem. 164, 375-381, 1987
A>Title: Identification and sequencing of cDNA clones for the rodent negative acute-phase
A:Reference number: A27199; PMID:87190405; PMID:2436907
A:Accession: B27199
A:Molecule type: mRNA
A:Residues: 1416-1483,'N',1485-1487 <SCH>
A:Cross-references: GB:M28297
A:Experimental source: liver, clone p13cDNA3
C:Genetics: 29/2, 90/3, 144/1
A:Introns: 29/2, 90/3, 144/1
C:Superfamily: alpha-2-macroglobulin
C:Keywords: proteinase inhibitor; thiolester bond
F:1-24/Domain: signal sequence #status predicted <SIG>
F:35-1487/Product: alpha-1 inhibitor III variant 1 #status predicted <MAT>
F:648-723/Region: bait region #status predicted
F:985-988/Cross-link: thiolester (Cys-Gln) #status predicted

```

Query Match

19.3%; Score 1417; DB 2; Length 1487;

Best Local Similarity 29.0%; Pred. No. 4,5e-72; Matches 454; Conservative 244; Mismatches 576; Indels 290; Gaps 55;

13 LCVCATLAAVAPGPRFLVTAAPGIRPGN---VTGVALLHRCPSQVTKALKLTAEN 68  
 9 LCLFSLALFLPFPASLL-----NGSKMTLVLPSQLYTETPEKCLHLTHLNEYTT 59  
 69 LTVSVLEAG-----VEKGSFKTL--TLPSLPLNSADEIYELRVGTODELLFSN 118  
 60 VTASILSQGTRKRLPDELAVDDLFHCLSFTHLPRLPSSEESLDINIGAHK---PFE 116  
 119 STRLSETRISVFIOTDKALKPKQGVKFRIVTLFSDPKPKYKTSLNIL-----IKPKS 173  
 117 RRVLVKNKESVVFQTDKPVKPGQSVKFRVVMKLNHP-----LNEFLPLAYIEDPKM 172  
 174 NLIQGLSQOSDLGVSITFOLSHPILDMSI---QVAVDQITVYQSFQSEVYLPRF 229  
 173 NRIQMODIKTEGKQLSFSLSAEPIQGVYKVIILKQGVKEE--HSFTMEFLPRF 229  
 230 EYTLQTPLYCSMNSKHLNGITITAKTYGKPVKQDVLTFL--PLSFWGKKINITTKPKI 286  
 230 GVDVKNPNAISVDELIINTVACALYTGKPVGHVAKISLCHGNP--SFSETSACKE--- 285  
 287 NSANFSFNDKMKVMSNGLS-----EYLDSSPGPVEILTTFESVTG 333  
 286 -----EDSELDN-----NGCSTQEVNITEFOLKENYILKMHQ--AFHVNATVEEGTG 330  
 334 ISRNSTVNFQKHODVIEFPDYTVLKRSLNPATVKTTRADGNQTLLEERNVITV 393  
 331 SEPSGGRLEVERTRNKKPLFLKADSHFRHGIPFVKIRLVDKGIDPNEQ---VFKA 386  
 394 TORNYEYSGNSGNOQKAEAVOKINYVPOSGTFIEFPILEDSESLDKAYPLGSKSS 453  
 387 QGAGYTNATTTQHGHLAK-----FSDISSISGYS-LNIKTYH--KESS 427  
 454 MAVHS-----LFKPSKTYIQLKTR-----DENIKVGSPPFEL--VVSQN 490  
 428 SCHHSCTARRHAEBHTAVAVYSLSKSYIYLDTEAGVPCNQIHTVQAHFLIKQGVLCV 487  
 491 KKLKELSYNVVSRGQLVANGK-----QNSTMSLTPENSWTKACIVYIYID 538  
 488 LPOIVFHYLVMAQGSILQTNHHTHQPVSQVQGFALPIPEFEMVAVAKLITITIP 547  
 539 DEEITSDVKIPQVLFKAKIKLYMSKVAEBSKYSLSISTVQPSIYIGVAVDSVNL 598  
 548 DEEVLADSTVFOVEKCLRNKHLSPSSQSLPASQTHMKTAS--PQSLGGLRAVDOSVLL 606  
 599 MNASNDITMENVHEL-----EL--YNYGYVLGMF 626  
 607 LKPEALFS--PSLIYDLPGMQDSNFISSVHPFEDEVDCLMYQPRDTEELTVGPY--GRE 663  
 627 NMSFAVQEGCLMVLTDANL---TKY-IDGYDNAEYERMRNEBEGHIVDIHDSLGS 682  
 664 KQVYRVRVMGLTAFNLKIKHPTCYEKMMVVLSPAVESLSPPRGSE--EMPLGVAK 722  
 683 SP-----HVKRHPETMIMLDTMAGRIYQEFVTVPDSTISVANG 724  
 723 SPKPKPPKDPKPPKQVETITRNHPETMIMDLTVNSSGTEVEMTVPDITTEKAKA 782  
 725 FVISEDGLGLTTTPELQAFPPFLNLPSVSVINGEFALETITFNLYKATEVKVIL 784  
 783 LCLSNDTGGLGSV--ATLQAFQFPFVELIMPYSVINGEAPMLKATVMYTLPSLPAVQL 841  
 785 EKSDKDIIMTSE-----INATG-HQOQLVLPSEGAIVLPFI-----RPHIGEIRI 832  
 842 EASPDPTAVPVGDDOSYCLGANGRHTSSMLVTPKSLGVNNSVSVEAOOSSELCSSQVA 901  
 833 TTTALSPASDAYTQMLIVKAGIEKYSOSILDLTDNRLOSTLTKLSFSPPTNVTS 892  
 902 TVPENG--KQDLYVKKVLIPEPKIKENHFPSSILCSDELFS---ETLSLLDPVYKDS 956  
 893 ERVQITAGDVLGSPINGLASLIRMPYGCGEQNMINFAPNIYILDTYTKKQVLTDLK 952

Db 957 ARAHFSVNGDILSIAIKNTQNLQMPYGCGEQNMVLPAPNIYLVKYNLTQOLTEKISK 1016  
 QY 953 ALSFMRQCYORELLAYOREDSFSFAFNVD--PSGSTMLSAFVLCFLEADYDIDQVNL 1010  
 Db 1017 ALGYIRAGYQREELNKKHDSYSAFBDHNGGCGNTMLITALVLSFAQARA.FIDESHI 1076  
 QY 1011 HRTYTMKGHQKNGEFPGRVISHSELQNGKSPVTLTAVITVSLGKRYKOPNID--V 1068  
 Db 1077 TDAFTWLSKQKQKDSGCFRSGSLFRNNAMKGVDEITLISAVITMALL--ESSLPDTPV 1134  
 QY 1069 QESHFLES-----EFSRGISDNTTALITLALSVS--PAKELANMLTRAEQGMQ 1122  
 Db 1135 SKALGCLASWETIEQGNNGSFVYTKLMAYAFALAGQERKRETLKSLDREAIRDMSI 1194  
 QY 1123 FM-----VSESKLSDSQPRSLDIFVAAVALSHFLQFQTSBGP-----IMR 1166  
 Db 1195 HMERQKTKSEGYLYTP-QASSAEVESAAYVALRL---TAQAPSEPDALSMGTIK 1249  
 QY 1167 WLSQRNSLAGPASTODITVALKALSEF-AALMNTERTNIQVTVGPSSPPLAVQ--- 1222  
 Db 1250 WLTQQNSHGPFSSQDTVVALDLASKYGAITFSKQKPLVTIQQSSGFSQKQVQVDSN 1309  
 QY 1223 -----PMAVINSANGFPAICQLVNYYNKAAGSSRRRRSIQNEAEDLVAV 1270  
 Db 1310 RLTLQVSLPDI PGNYTVSVSEGCVVAQTILRYMPLR-----KQAPAPALKVQT 1360  
 QY 1271 K-----ENKDDLNHDVNLVCTSPSGP-GRSGMALREVNLSGFVWPSAISLSEYTKVE- 1324  
 Db 1361 VPLTNKNGQNSFOISLEISITGSRPASNVIVADVKMLSGF-----IPKPTVKLER 1414  
 QY 1325 YDH-----GKNTLYDSVNETQPCVNIPIAVBNFKVNSNTODASVSIYDVEPRQAVR 1376  
 Db 1415 LEHVSREVTNNVNLVLYDQYTNQTLSPFSFIQQDIPKKNLQPMIVKYYDYREDEVAF 1474  
 QY 1377 SYNS 1380  
 Db 1475 EYSS 1478

RESULT 9  
 A26122  
 alpha-2-macroglobulin precursor [validated] - rat  
 N:Alternate names: alpha-2M  
 C:Species: Rattus norvegicus (Norway rat)  
 C:Date: 16-Aug-1988 #sequence revision 16-Aug-1988 #text change 31-Dec-2000  
 C:Accession: A26122; A05278; S03431; S33843; A26141; I52403; I54013  
 R:Gehring, M.R.; Shleis, B.R.; Northemann, W.; de Bruijn, M.H.L.; Kan, C.C.; Chain, A.C.  
 J. Biol. Chem. 262, 446-454, 1987  
 A:Title: Sequence of rat liver alpha-2-macroglobulin and acute phase control of its mess  
 A:Reference number: A26122; MUID:87083488; PMID:2432068  
 A:Accession: A26122  
 A:Molecule type: mRNA  
 A:Residues: 1-1472 <GB>  
 A:Cross-references: GB:J02635; NID:G202591; PIDN:AAA40636.1; PID:G202592  
 J.Hayashida, K.; Okubo, H.; Noguchi, M.; Yoshida, H.; Kangawa, K.; Matsuo, H.; Sakaki, Y  
 J. Biol. Chem. 260, 14224-14229, 1985  
 A:Title: Molecular cloning of DNA complementary to rat alpha-2-macroglobulin mRNA.  
 A:Reference number: A05278; MUID:86033908; PMID:2414291  
 A:Accession: A05278  
 A:Molecule type: mRNA  
 A:Residues: 178-227/420-489, 'L', 491-526 <HA>  
 R.Kunz, D.; Zimmermann, R.; Heisig, M.; Heinrich, P.C.  
 Nucleic Acids Res. 17, 1121-1138, 1989  
 A:Title: Identification of the promoter sequences involved in the interleukin-6 dependen  
 A:Reference number: S03431; MUID:89160243; PMID:2466233  
 A:Accession: S03431  
 A:Molecule type: DNA  
 A:Residues: 1-22, 'V', 24-164 <XU>  
 A:Cross-references: EMBL:X13983  
 A:Note: The authors translated the codon GTG for residue 23 as Leu  
 R.Watburton, M.J.; Coles, B.; Dundas, S.R.; Gusterson, B.A.; O'Hare, M.J.  
 Eur. J. Biochem. 214, 803-809, 1993  
 A:Title: Hydrocortisone induces the synthesis of alpha(2)-macroglobulin by rat mammary m





Query Match	19.1%	Score 1402	DB 2	Length 1451
Best Local Similarity	27.7%	Pred. No. 3.1e-71		
Matches 439	Conservative 250	Mismatches 527	Indels 370	Gaps 54

QY 680 -----LSSPH-----VRKPEPTWIMLDTMGRIYQEP 710  
Db 698 SSEVSGTLQKSSRRPPEPPEDPPKPLAEIRKPIPETWMDIYVNSTGVAEVE 757  
QY 711 VTPDPSITSWAATGVISEDGLGTLTTPVELQAFQPFILNLPSVINGEFALEITI 770  
Db 758 MTVPDITTEWKAGALCLISNDTGLSSV-VPLQAFQPFVSVSVVGAEPMLKATV 816  
QY 771 ENYIKADTEVKVILIEKDKDILMTSE-----INATG-HQQTLLVPSEDAIVLPFIR 824  
Db 817 MNVLPTSMRMSVCLSEASPDFTAVPVGDDHSDYCLISANGRTSSMLV-----TP 864  
QY 825 THLGEIPIITVATLSPAS-----DAVQMIIVKAEIGEKSSQSTILLD 868  
Db 865 KSLGNVAFSVSVARQSSPECCSEVATVPATGKTDIVYKVLIVPEPCIKQEHFNLSIFCA 924  
QY 869 TDNRLQSTLKLTSFSPFPNTVTGSEVQVLTALGDVLPSSINGLASLIRMPYGCGEQNMN 928  
Db 925 SDAEIS---EKMSLIVPEPTVVKDSARAHFVSMGDILLISAIKNTQNLHMPYGCGEQNMV 981  
QY 929 PAPPIIYLDLITLTKKQITDNLKEKALSPMGQYQRELLYRREGSSAFG--NYDPSGCT 986  
Db 982 PAPPIYVLKTLIDTQOLOTOKIKTKALGFLNAGQREILMYGHKGGSYAFPDQGEERBGN 1041  
QY 987 WLSAFLRCLAEADPYDIDONVLAHRTYTLKCHQKNGSEFMFGRIHSELQGNKSPV 1046  
Db 1042 WLTPFLVKSQAQAPAFIDESHITIAFTLSSQQKDNCGFRSSGSLFPHDI---KHPV 1097  
QY 1047 TLTAIVITSLIGYKRYQPNIDVQESIHFLES-----EFSRGISDNTYALITYALSSVGS 1101  
Db 1098 -----VSKALSCLESSMKTEQGRNANFVYTKALMAVAFALAGN 1136  
QY 1102 -PKAKELANLWTRABEGGMQVW-----SSEKLSDSWQRRSLDIEVAATALLSHFL 1154  
Db 1137 QDKRNEELIKSLDEBAIKEDNSIHMERPOKREKKEHNLKYP-QASSIEVENNAVVLART- 1194  
QY 1155 QFOTSEGI-----IMRWLSROBSNLGFASTQDPTVTLKALSEFALAMNERTN 1204  
Db 1195 ---TAQAPRPEDLITLSRSTIMTLTKQNSNGFSSDTQVVALDLASKGATVPSRRQ 1251  
QY 1205 IQ-VTVGPSSPSPPLAVQ-----PMAVNISANGFGFALCOLNVVVNVKA 1248  
Db 1252 TSLVITIGSTGSPQKFEVENSNCILLQOVLPDPIPGDYTISVSGEGCVAAQTLLRINMHL 1311  
QY 1249 SGSSRRRRRSIQNGEAPFLDVAVK---ENKODLNHYDLAVCTSPSGP-GRSGMALHEVNL 1303  
Db 1312 E-----KOOSAFALRVQTVPLTCNNPKGNSFOISLEISYTSGRPASNNVIADVKA 1362  
QY 1304 LSGFMVPSSEALISSEYTKV-----FYDHGKLNLYLDSYENETQFCVNIIPAVNPFKY 1354  
Db 1363 LSGF-----IPLKPYKLEERLEHISRTIEVSNNVLLYDQVTTNOTLAFSPITQODISV 1416  
QY 1355 SNTQDASVSIVDYEPERRQAVRSYNS 1380  
Db 1417 RNLPQPAIVKVVYDYETDEVAVAYESS 1442

RESULT 11  
S13495  
Pregnancy zone protein - human  
C:Species: Homo sapiens (man)  
C:Date: 17-Jul-1992 #sequence\_revision 17-Jul-1992 #text\_change 16-Jul-1999  
C:Accession: S13495; S08106; A34230; P80077; S21637  
R:Deviant: K.; van den Berghe, H.; Caselman, J.J.; Marynen, P.  
Biochim. Biophys. Acta 1088, 95-103, 1991  
A:Title: Primary structure of pregnancy zone protein.  
A:Reference number: S13495; PMID:9113734; PMID:189698  
A:Accession: S13495  
A:Status: not compared with conceptual translation  
A:Molecule type: mRNA  
A:Residues: 1-1482 <DEV>  
A:Cross-references: EMBL:X54380; NID:G35824; PIDN:CAA38255.1; PID:G35825

FEBS Lett. 262, 349-352, 1990  
A:Title: A genetic polymorphism in a functional domain of human pregnancy zone protein:  
n.

A:Reference number: S09106; MUID:90242963; PMID:1692292

A:Accession: S09106

A:Molecule type: DNA

A:Residues: 668-690, 'M', 692-753 <MAR>

A:Cross-references: EMBL:X51541

R:Scitrip-Jensen, L.; Sand, O.; Kristensen, L.; Fey, G.H.

J. Biol. Chem. 264, 15781-15789, 1989

A:Title: The alpha-macroglobulin bait region. Sequence diversity and localization of cle

A:Reference number: A34230; MUID:85380162; PMID:2476433

A:Accession: A34230

A:Molecule type: protein

A:Residues: 670-752, 'Q', 754-759 <SOT>

C:Genetics:

A:Gene: GDB:P2P

A:Cross-references: GDB:120330; OMIM:176420

A:Map position: 12p13-12p12.2

A:introns: 73/3; 116/2; 146/3; 169/3; 703/1; 753/2

A>Note: the list of introns may be incomplete

C:Superfamily: alpha-2-macroglobulin

F:685-735/Region: bait region

Query Match 19.0%; Score 1396.5; DB 2; Length 1482;  
Best local similarity 28.3%; Pred. No. 6.5e-71;  
Matches 432; Conservative 280; Mismatches 587; Indels 225; Gaps 53;

12 LHCCTALAAVAPGRFVLTAPGIIIRPGANTIGVLEHPCSOVYTKAEILKASNLTV 71

14 LLLLSASDSNSTEQYVAVLPSLHTEAPKK-GCVLLSHNETVYVASLESGRERS- 71

72 SVLEAGVFEKSGPK--TLTLPPLNSADEIYELVGRTODELFSNRLSFEFKRI 129

72 --LFTDLVAKDLFRCHVSFTLRISASSEVAFSLIQIKGPTOD--FRKRVVLTNOSTL 127

130 SVFIQTDAKLYKQKQEVKRIIVTLFSDPKPYKTSNLTL-IDPKSNLLOQWLSQOSDLGV 188

128 -VFQTDKPMVKPGQTVAFRVVSDENFRPNELIPLILENPRNRJAQWQSLKEAGI 186

189 ISKTFQLSSHPILGDMISIQVANDQTYQ-SFOVSEVYLPRFEVTLQTPLYCSMNSKHLN 247

187 NQLSFPLSSSEPIQOSYRVVQTESGRIQHPFVEEFVLPKFEKVQVPKIISIMDEKVN 246

248 GTIRAKYVGPVKGDLVTLFLPSFMGKKNITKTFKINSANPSFNDEEMKRVMSNN 307

247 ITVCGEYTKGKVEGLAT--VSLCKLSRVLANCDQKQVCEBPS--QQL-----NSN 293

308 GL-----SEYLDLSSPG--PVELLTVTESVTGI--SHNVS--TNVFPKQHDYIIEF 353

294 GCITQOVHTKMLQITNNGFEKMLRVEARIREGTDLEVTANRISBITNIVSK-----LKF 348

354 FDYTVTLKPSLNFATVATVADGNQLTLERRNNVITVTOKNYTEWSSGNSG----- 408

349 VKVNSHFQGGPIFPA--QVLVDGKGVIPNKL--PFISVNDANYYSNATNEGLQFS 404

409 -NQMKAAYQKI--NYTVPGSTFKIEPILDESSBLQKAYFISGKSSMAVHSLFKSPS 464

405 INTTISISVNLKLFVRVFTVHPNLCFRYSW--VAEDHQAO-----HIANRVSFL--S 451

465 KTYIQLKTRDENIKV-----SPPELVNSGKKLKELS--YNVVSRQGLVAVG----- 510

452 GSYHLEPVAGTGLDGHETITTAHTLNROMGSELSELSFHYLMAGVIVRSSTHLLPV 511

511 ----KQNSTMFSLPENSMTPRACVIVYIEDGSEIISDLKIPQVLVFNKKIKLYNSKV 566

512 ESGDMKGSFALSPVESDVADIAEMFIPAILPDESVGVDSSEKFEIENCLANKVDLSFSPA 571

567 KAEPSEKSLAISTQPSIYIGVIAVDSVNLMAANDIMENYVHLE----- 615

572 QSPASHAHLOVAA-POSTCALRAVDQSVLLMKPEALSVSSVYNLLTVKDLTNPDNV 630

616 -----LVNTGYLLGMFNMS---FAVFGCGMLVLTDAVLTKDYIDGVYDN 657

Db 631 DOOEEOGHCPRPFFIHNCAIYVPLSSNEADISFLKGMKLKFTMSKIRPKSCSVIPS 690

Qy 658 AE-----YAERFMENEG--HIVDI-IDFSLGSSGP-VRKHPETMIWLDT 699

Db 691 VSAAGVGGYAGGLGVERPYPQIGTYVATPLANEQSGPVPEVTRSYFPEWTIMELV 750

Qy 700 NMGRYIOEPEVTPDSITSMVATGVVISSEDLGLTTTPVELQAPOPFPILNLPYSVI 759

Db 751 AVNSGVAEGVAVPPITTEWKAGAFCLSDALGISST-ASLRAPQFFVELTWPEYSVI 809

Qy 760 RGEPALEITFNYLKDATEVKYIIEKSKFDILMTSSSEINATGHQCTLVPSDEGATVL 819

Db 810 RGEVFTLKATVNLVLPKICIRVSVQLKASPAF-----LASQNTGSESSYCQGE-RQLTS 863

Qy 820 FPIRPHTEGIPITVLAASPTA-----SDAVOMLVVAKGIEKYSOS 863

Db 864 WTVPTKLGNVNSVSAHEAQSLQEGNEVEVEPEIKRQDTYKTLVLEABGIEQKITS 923

Qy 864 ILDLTDNRLQSTLKTLSFSPPTVTSERVQITAGDVLGSPINGLASLRMPYCGGE 923

Db 924 ---SMTCASGANVSEQLSLKLPENVVKESARASFVLDGLISAMQNIQVLMQPGCGE 980

Qy 924 QNMVNPANITYIIDYITKKQQLTDNLKELASFRQGYQRELLYQREDSFSAFG-NYDP 982

Db 981 QNMVLPANITYVNLNYNETQOLQOIKAVAGYLITGYQRLNKKHQDGSYTFGERYGR 1040

Qy 983 S-GSTMASAPVLCPLADADYIDIDONVLRHTYTMKGHQKSGEPMDCRVHSELOGG 1041

Db 1041 NQNTMTLAFVLTQFQAPASVIFIDENHITQSLTWLQMOQKGCRRSGSLNNAIKGG 1100

Qy 1042 NKSPVTLTAVTSLGKRYQPNIDVQSHPLS-----EFSRGISDNTYLTALITVA 1095

Db 1101 VEDEATLSAVITLALIEIPVTPNPIVRNALPCLSEAMVAKGCTHG-SIVYTKALLAVA 1159

Qy 1096 LSSVGP-KAKELNMLTWAEQEGMQFVSSSEKLSD-----SNQPSLDIEVAAYAL 1149

Db 1160 FSLGKQNGQREITLISLDEKAVEDNLVHERPQRKAPGHLYQFOASAEVEMTSYVL 1219

Qy 1150 LSHFLOQFQSEGIP-----IMRWLSHQNSISGSPASTODTVALKAUSEF--ALL 1197

Db 1220 LAVL---TAQAPATSGDLTSATNIVKIMKQNAAGFSSTODTVALHALLASRYGAATP 1275

Qy 1198 MTERNTNIQVTVGSPSPSP-----LAVQPMVAVN-----ISANGFPAICQLNV 1242

Db 1276 TREKT-AQVTVGDSDQTFSTNFOVDNNMLLQOISLPELPGRYVITVGERCVYIQTSM 1334

Qy 1243 VYNKASGSSRRRSIQNDPAFDVAVXENKODLN--HYDLNVCTSPSGPGR--SGMA 1297

Db 1335 KYNILPE-----KEDSPFALKVQTVPQTCDGKHAHTSFQISLTISYTGNRPASNMV 1385

Qy 1298 LMEVNLISGFMVSEALIS---SEYKAYEYDHGKLNLYLDSVNEFOFCNIPAVANPKY 1354

Db 1386 IVDVKNVSGFTPLKFTVKMLERSSVSREVSNNHVLIVEOVTNOTLSFSEWVLDIDIV 1445

Qy 1355 SNTQDASVSIVDYVPEPRQAVRSY 1378

Db 1446 GDLKPAIVKYYDYETDESVAEY 1469

## RESULT 12

A41185 alpha-2 macroglobulin M0G1 - mouse

C:Species: Mus musculus (house mouse)

C:Date: 28-May-1992 #sequence\_revision 28-May-1992 #text change 12-Apr-1995

C:Accession: A41185

R:Overberg, L.; Torrekens, S.; Van Leuven, F.; Van den Bergh, H.

J. Biol. Chem. 266, 16903-16910, 1991

A:Title: Molecular characterization of the murinoglobulins.

A:Reference number: A41185; MUID:91358495; PMID:1840592

A:Accession: A41185

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-1476 <OVR>  
 A:Cross-references: GB:M65736  
 C:Superfamily: alpha-2-macroglobulin

Query Match 19.0%; Score 1395.5; DB 2; Length 1476;  
 Best Local Similarity 27.7%; Pred. No. 7.4e-71;  
 Matches 444; Conservative 269; Mismatches 555; Indels 311; Gaps 50;

Query 13 LCVCTAALA-----VAPRFLVTAPGIIIRPGNVITGVELLEHCPQVTKAELKLT 65  
 Db 9 LCLFSLTLLAFHSASLNDGSKMVLVPS-----QLYTTPKICHLAYQANE 56  
 Qy 66 ASNLIVSVLEAG-----VPEKSPKTLT--LPSLPLNSADEIYELVTGRTQDEL 115  
 Db 57 TTVTVAISLVSGGRKNLFDLVLVDKOLFQVSGFIILRLSSDEDEFLVDIKQPTHE--- 113  
 Qy 116 FSNSTRLSFETKRISVFIOTDKALYKPKQKQKRYITLESDEKPYKTSINT--LIDPKSN 174  
 Db 114 FSKRAVALVKNKESVVFQTDKPVYKPGQSVKFRVSMKMLRPLNELPLVYIEDPKKN 173  
 Qy 175 LIQOMLSQSDLGVIKSTFQLSHPIIDGMSIQV--QVNDQTYQSFOVEEYVLPKPEVTL 233  
 Db 174 RIMQMRDITENGILKQMSFSLAEPICQFYKIVAKHSEKSEKSHSTVMEFVLPRFNDVL 233  
 Qy 234 QTPLYCSMNSKHLNGITITAKYTYGKPEKQDVTLTFLPLSPWQKKNITTYFKINGSANFS 293  
 Db 234 KVPNAMSVDDEVLSVACGKTYGKVPGHV-----KINVCERE 273  
 Qy 294 FNDEKKNVMSNGIS---EYLDLSPGP-----VEILTTYESVTGI--SRNVSTN 341  
 Db 274 TQCREVNSQLD--NNGSGTOEVNITELQSKKRYEVOLFHNATVIEEGGLSEFSSGTYK 332  
 Qy 342 VFFKQHDYIIEPFDYTVLKPSLNPTATYKVRADGNOLTLERRANNVITATQRYTEY 401  
 Db 333 IERTYKLI--FLKADSHRHGIPFVYKRLVDIKCDPIPNK---VFIKAQELSYTA 386  
 Qy 402 WSGNSNGQKMAVOKINYTPQSGTFKIEFPLEDSSELQKA-----YFLGS 450  
 Db 387 TTTDQHG-----LAEFSIDTTCISGS--LHVKVNHKEDESGSYFCMBE 429  
 Qy 451 KSMNVHSLFK--SPSKTYIQKTRDENT-----KVSGPELVVSGNR--LKEI--SY 498  
 Db 430 RHASAGHVAAYVSLSKSYIYLDTESSILPCNQIHTVOAHF--LIKGLDGLVKELEIFY 487  
 Qy 499 MYVSRQGLVAVGKQNSTM-----FSITPNSMTPKACVIVVYIEDDELISDV 546  
 Db 488 LVMALHSLITQGTGHTHQVEPGEAPVKGKALETIPEFSSVPMKMLITYILPDGEVIAS 547  
 Qy 547 LKIPVQVLEKNIKILYWSKYKAPSEKVSRLISVTQPSIVGIVAVDKSVNLNANSNDIT 606  
 Db 548 VAPTEIKCLRNVDLRFSTQSRLPASQTRLQVTAS--PQSLCGIRAVDQSVLLKPESELS 606  
 Qy 607 MENNVHELELYNTGYLLGMPNNSFAVFOEGCMLTLDANLTKDYIDGVYDAAEYAE--- 662  
 Db 607 PS-----WYIYLP--GMQONKF-----VSSRLSEDOEDCIYSSWLAEKHTN 647  
 Qy 663 -----RPMEN-----BEHYIDH-----DELSGSPH----- 685  
 Db 648 LVPHGTEKDVYRVEDMGLTAFNTMLKPLIICFDGMPISAPRVEFLAETPEISWEL 707  
 Qy 686 -----VRKHPETWILDTNMGRYIQAQEPVTPDSITSM 720  
 Db 708 RTTLSKRPPEPRKPDSSNDPLTETIRKTFPEETWVDIYTVNSTGLAEVEMVDPDITTW 767  
 Qy 721 VATGFIISDGLGLTTTPVELQAFQPFPIPLNLPSYVIRGEFALEITIFNYLKDATEV 780  
 Db 768 KGALCTISNDTGIGLSSV--VPLQAFKPFVEVSLPSYVVRGEAFMKATVMYVLPFSQM 826  
 Qy 781 KYIIEKDFDLIMTSEINATGHQOTLVPSDEGATVLPPIRPHLIGELIPITVTLST 840  
 Db 827 SVQLEASPDFTAVPGVD-----QDSYCLISANGRTSSMLVTPKSLGWNVSVALEAQ 880  
 Qy 841 AS-----DAVTQMIIVKAGIEKYSQSILDLTNRLOSTLKTLSFSG 884

Db 881 SSEPGSEVATVPATGRKDTVAKVIVPEEIKQEHFTSSIFCASDAEIS---EXMSSGP 937  
 Qy 885 PENTVGSERVQITAGIVLGPISINGLASLIRMPYGGGEOMINPAPITVILDLTKKKO 944  
 Db 938 PPTVVKDARAHFSSWGDILSSALINTONPLHMPFGGCEQNNVLPAPITVILKINTEQ 997  
 Qy 945 LTDNLKERALSFMRQGYREILYQREDESPSFAFG--NYDPGSGTSLSAFVRCPEADPY 1002  
 Db 998 LTQKIKTALGRLRAGYRELINTYKHKQSYARFGQNERGNTWLTFLVLSFAQARAF 1057  
 Qy 1003 IDIDQNLHRTYTLKGHQKNGEFPDGRVHSELOCGKNSPTVLTAVYTSLSIGYKY 1062  
 Db 1058 IFIDESHITHAFWLSQKQKNGCFRSGSIFNNAMKGVDDEMTLSAYITMALLESLP 1117  
 Qy 1063 QPNIDVGSIHFLS-----EFSRGIQNTYLTALITYLSSVSGSP--KAKEALNLTWRAE 1116  
 Db 1118 ATHPVASKALSCLESQWKTIEQERNASVYTKALMAVAFALAGQNKREDELKSLDEBAI 1177  
 Qy 1117 QEGGWQFNV--SSESLKSDSW-----QPRSLDIEVAAYALLSHFLQFOTSEGIP----- 1163  
 Db 1178 KENNIIHKRPQKSKSHHLYKQASAEVMAHYVLAAL-----TQAPAPSPEDLTLS 1233  
 Qy 1164 --IKMWLSGRNSLGGFASDTDTVVALKALSEFAL--NMTERNTIQTVTYGPSSP--- 1215  
 Db 1234 MSTIMWLTQKQNSNGFSTQDTVVALDALSKYGAVTFSRQKTLVTIQSTGFSQKQ 1293  
 Qy 1216 ---SPLAVQPMAY-----NISANGRFALICQNLVYVYVYKAGSSRRRSIQNEAFD 1265  
 Db 1294 VENSRRLLQOVALLPDIPGDYTIISVSGCYVAAQTMRLYNNHLE-----KOLSAFA 1344  
 Qy 1266 LDVAVK-----ENKQDLNHDVLDVNCVSPSGP--GRSGMALMEYNLISGFVWPSBAISLSRTV 1320  
 Db 1345 IWQVTVPLTCNNPKGHSFQSLSTYSGSPASNMVYADVKMLSGF-----IPKKTV 1398  
 Qy 1321 KKV-----EXDHGKLNLDSVNETQFCVNIPAVRNFKVNSTODASVSIYDYEP 1371  
 Db 1399 KKLRLRHVSRFVSNVNLVLYLDQVTWQTLAFSPFIQODIPVNLQAVIYKVVDYETD 1458  
 Qy 1372 ROAVNSYNS 1380  
 Db 1459 EWAFAYSS 1467

RESULT 13  
 A20872  
 ovostatin precursor - chicken  
 N:Alternate names: ovomacroglobulin  
 C:Species: Gallus gallus (chicken)  
 C:Date: 05-Jun-1987 #sequence revision 13-Mar-1997 #text\_change 31-Mar-2000  
 C:Accession: I50671; I50672; S29836; A20872; A33715  
 R:Nielsen, K.L.; Soltup-Jensen, L.; Nagase, H.; Thøgersen, H.C.; Ezerodt, M.  
 DNA Seq. 5, 111-119, 1994  
 A:Title: Amino acid sequence of hen ovomacroglobulin (ovostatin) deduced from cloned cDN  
 A:Reference number: I50671; MUID:95218210; PMID:7535598  
 A:Accession: I50671  
 A:Status: translated from GB/EMBL/DBJ  
 A:Molecule type: mRNA  
 A:Residues: 1-1473 <NIE>  
 A:Cross-references: EMBL:X78801; NID:G671863; PIDN:CAA5384.1; PID:G671864  
 A:Accession: I50672  
 A:Status: translated from GB/EMBL/DBJ  
 A:Molecule type: mRNA  
 A:Residues: 20-1473 <N12>  
 A:Cross-references: EMBL:X78801; NID:G671863; PIDN:CAA5384.1; PID:G671865  
 R:Nielsen, K.L.; Soltup-Jensen, L.  
 Biochim. Biophys. Acta 1162, 230-232, 1993  
 A:Title: Evidence from sequence analysis that hen egg-white ovomacroglobulin (ovostatin)  
 A:Reference number: S29836; MUID:93192299; PMID:7680377  
 A:Accession: S29836  
 A:Molecule type: Protein  
 A:Residues: 977-980, 'N', 981-1007, 'M', 1009, 'K', 1010-1028 <N13>  
 R:Nagase, H.; Harris Jr., E.D.; Woessner Jr., J.F.; Brew, K.

J. Biol. Chem. 258, 7481-7489, 1983  
 A:Title: Ovostatin, a novel proteinase inhibitor from chicken egg white. I. Purification  
 A:Reference number: A92427; MUID:83238315; PMID:6408074  
 A:Accession: A20872  
 A:Molecule type: protein  
 A:Residues: 37-49 <NAG>  
 R:Englind, J.J.; Salvesen, G.; Brew, K.; Nagase, H.  
 J. Biol. Chem. 264, 8779-8785, 1989  
 A:Title: Interaction of human rheumatoid synovial collagenase (matrix metalloproteinase kinetics and identification of matrix metalloproteinase cleavage sites.  
 A:Reference number: A33715; MUID:89255343; PMID:2470748  
 A:Accession: A33715  
 A:Molecule type: Protein  
 A:Residues: 710-717, 'I', 'T', 'A', '723-742', 'L' <ENG>  
 C:Superfamily: alpha-2-macroglobulin  
 C:Keywords: alternative initiators; egg white; proteinase inhibitor

Query Match 19.0%; Score 1395; DB 2; Length 1473;  
 Best Local Similarity 27.0%; Pred. No. 7.8e-71;  
 Matches 421; Conservative 309; Mismatches 522; Indels 308; Gaps 59;

```

18 AALANAPGRFLVTPAGIIRGGANTIGVELLEHCPQVTVKAEILKTSMLTVS---VL 74
   ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 33 AAAGKEPPQYVLMVPAVLQ-----SSPSQVCLQFENL---NQTLSRVVL 76

75 EAE-----GVFEKSGFKTLPLSL-----PLNSADEIV-ELRVTGRTO-----ELFENS 119
   ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 77 EYDINTNTIFEKNTTNSGLQCLNFMIPVTSVSLAFISFPAKGTTPDKRRSRVMIMM 136

120 TRLSFEKTRISVFIOTKALYKPKQEVKFRIVLTFSDEPKPKYKTSMLNI--IDPKSNLIQ 178
   ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 137 E--SF-----VFVQTDKPIYKPGQSVMPFRVALDFNFKPVQEMKPLIAYVDDPQNNRIQ 188

179 WLSQSDGLGVISKPEQLSHPLGDMWSIQV--QVNDQTYGSGFOYSEYVLPRFEVTLQPL 237
   ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 189 WQNTSEINIVQIEPLLEBPILGNKYIIVTKSGERSHSFLVEEYVLPFEDVITVAPG 248

238 YCSMNSKILNGLITAKTYGKPKVKGDVLTPL-PLSFWGKKK-----NITKTKINGSA 290
   ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 249 SLTWSDSLVYKICAVITYGQVBEKGQVLSVCRDSDSGRCKSPVCGSPFKDLDTQCL 308

291 NFSFND--EMKNVMDSSNGLSSEYILSSPGFVEILITVTESVTGISRNVSTNVPKQHD 348
   ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 309 SHILSSKVFELNRI-----GYKRND-----VKAIVTEKQVCMLTATQSI--TQVM 355

349 YIIREFDTYVLPKPSLNTATYKTRADQNLTEERNNNVITVQNTIEYVSGNSG 408
   ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 356 SSLQFENDVHHYRRGIPFGQIKLVDKDQNSPI-----SNKVIQLFVNKNKTH----- 402

409 NQKMEAVQKINYTPQSG--TFKIEFPILDESSSELQKAYFLGS-----KS 452
   ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 403 -----NFTDINDIAPFSIDTSKIPD-PELSLKALKYKTDQCHSEGMIBPSYDA 451

453 SMAVSLFKSPSKTYIQLKTRDENIKVSP-----FELVVSQNKRLKELS--YMWVSRG 504
   ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 452 SLVQORLY-SWTSSFRVLEPLMKDMSCGQKEMITVYIYILNTEGEYHINIYVFFYGMAGK 510

505 QLVAVGK-----QNST-WFSLTPENSMPKACVITYYIEEDDEIISDVLKIPVQVLF 555
   ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 511 KIVLGEIKVNIQADQNGTFMPLVNVNEKMAPALRLVYMLHPAKELVADSVRSIEKCF 570

556 KNIKILVYSKVAEPESEKSLRISVTPDSIVGIYAVDVSNVLMANASNDIMENV--VHE 613
   ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 571 KNIKVLQSEKQMLTTSNVSLVIEAA-ANSCFAYRAVDKSKMLLSEKLSAEITTYILHP 629

614 L-----ELVNTGYLL-----GMFNSFAVPOECGLMVLTD 643
   ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 630 IODLGVIYFNGNLNLEDDPQDPCVSSDI.FHKGLYRPLTSGLGPDVYQFLRDMGKKEFTN 689

644 ANL-----TQDYI-----DGYDNAEYAEAFMEENEGHIYDIDHDFSLGS 682
   ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 690 SKIROPVTCRETVRPPSPFLNAGFTASTHHVKLSAEVARE--BRGKHILDT----- 740
  
```

```

683 SPVHKHFPETWIM---LDTNMGRYIOEFEEVTPVPSITSWATGVIISDLGLTTP 739
   ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 741 ---IREFFPETWIMDILINSTG---KASVYIPIPTITWKMKASAFCEVLEAGFGM-SVP 793

740 VELQAPPEPFIPLNPSVIRGEEFALETIENYILADATEVVKYIIEKSDKFDLMTSSEI 799
   ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 794 ATTLAPQPFVVDLTLPPSIILHGDFLVRAVAVFYLNHCINVL-----L 839

800 NATCHQOTLLVPSDEGA-----TVLPFIRPHLGEIPITYVA-----L 837
   ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 840 ESDYQAKLSPEDDGVCVKIKRSYWNIFPKGTGVLFSITAFETNDACEEALRNI 899

838 SPASDAVQMILVKAEGIEKYSQSILDLDTNRLQSTLKTLSFSPFPNTVTSERVQI 897
   ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 900 RIDRDTQIRALVPEPGIRRETONFLCMKDVIS---QDAVIDLPTNVVSGSPRSF 956

898 TALGDVIGPSINGLASLRMPYCGGROMINFAPNYIILDYLRKKQLTNNLEKLSFM 957
   ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 957 SVGDIWGTALQVNHQLQMPFGNQQNVYLFAPNIVLDYLDKTRQLSDEVSKTITGYL 1016

958 ROGQRELYQREDEGSFSAFGNTPSGSTWLSAFVLRCPLEADPYIDIDQNVLAHRTYWL 1017
   ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1017 VSGYQKQLSTKHPDGSITGTGIRDKGNVTLTAFVYKSPFAASRPFIYIDNVQAQITLWL 1076

1018 KGHQKNSGEFMDPGRVHSELQGNKSPVTLTAYIVTSLGYRKYOPNIDVQESIHFLS 1077
   ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1077 ATKQKTGGCFQSTGILVNNMKGVENELSLSAVITITALLGASHMHTVIRNAPCLET 1136

1078 EFSRGISDNTALLITYALSVSVPKAKEN-LMMLTRAEQEGCMQPVSSF-----SK 1130
   ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1137 ASEKNITDIYQALVNAFAFLAKAEICSEFLBELQKSAEYVGSKRWENQGSAPKSH 1196

1131 LSDSWQPSRSDIEVAAYALS-----HFLQFQTSSEGIPIRMWLSRONSIGCFASDTDT 1185
   ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1197 LLDHVQ--STDVETISVLLALLKYKRSQEDLTKASAIYQWIIKQNSIGCFASMDTV 1254

1186 VALKALSEFAL-----MNTERTNIQV-TYTPGS---SPSLAVQPMAYNIS 1229
   ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1255 VALQALAAAYASVNSVQVNIKINSKPTPEKFTVNNERLLQLQTPPLQV--RGKYST 1313

1230 ANGRPAICQLNVNVYVYKASGSSRRRSIONQCAFDPDVAVKEN-----KDLNHDVLAN 1284
   ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1314 VNGTCVLIQTLARVNIHLP-----ECAFGRSLSVQTSNASCPSDQPKGFIVL 1362

1285 CTSPSGP-GRSGMALHEVNLISGFMYPSAISLSB-----TVKRYEDHDKMLYLDVSF 1337
   ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1363 ISSYTGKSSNNVITIDVKLSGF-VPKS-SLDQLIDHTVAVQEKKNHVLILYGNIL 1420

1338 ---NETQFCVNIPAVNFVSNTPDASVSIYDYEBRQAVRSYNSVYKLSSCDLCSDV 1393
   ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1421 QKRRKEVTFSEV---QDFVVTHTPKAPVQIYDYETEEVAVAEYMS-----LCRGV 1468
  
```

RESULT 14  
 A42210  
 alpha-1-macroglobulin precursor - rat  
 C:Species: Rattus norvegicus (Norway rat)  
 C:Date: 31-Dec-1993 #sequence\_revision 31-Dec-1993 #text\_change 16-Jul-1999  
 C:Accession: A42210; A26124  
 R:Warmegard, B.; Martin, N.; Johansson, S.  
 Biochemistry 31, 2346-2352, 1992  
 A:Title: cDNA cloning and sequencing of rat alpha 1-macroglobulin.  
 A:Reference number: A42210; MUID:92172859; PMID:1371696  
 A:Accession: A42210  
 A:Molecule type: mRNA  
 A:Residues: 1-1500 <WAR>  
 A:Cross-references: GB:M84000; GB:J05359; NID:9205383; PIDN:AAA1591.1; PID:9205384  
 A:Experimental source: strain Sprague-Dawley, liver  
 A>Note: sequence extracted from NCBI backbone (NCBI:87164)  
 R:Jonberg-Holm, K.; Reed, D.L.; Roberts, R.C.; Hebert, R.R.; Hillman, M.C.; Kutney, R.M.  
 J. Biol. Chem. 262, 438-445, 1987  
 A:Title: Three high molecular weight protease inhibitors of rat plasma. Isolation, charac  
 A:Reference number: A26124; MUID:87083487; PMID:2432067

D	b		15	LLLLPRDATTATGKPRVVLVPSBELYAGVPEKVCH-LMHLNETYTLANTLVETGVQXSNL	73
Q	y		70	TVSVLAEAGVFEKSGFKTLLTPSLPLNSADEIYELVTKGTODEILFSGNSTRLSFEFKRI	129
D	b		74	LID-----QAVDKOSSYCSSFTISRPL-SPSALIAVEIKGPHTHFI-----KKKSMITIKAE	124
Q	y		130	S-VYIOTDKALYRKQEKVEKRIYTLFSDPKPKTSLNLT-IKDPKSNLTQWLSQSDL-	166
D	b		125	SPVFOYDDEKRIYKPGTYKFRVAVSDISFRPNEFFVYVYIENPKRNIIFQW-QANDLP	182
Q	y		187	-GVISKTFOLSHPIILGWSIQVQND-QTYVQSFQVEEYVLPRKEVLTQPLCYMNSK	244
D	b		183	GGHQLSHFSLSVBPALGIYKVVQDQSGKKIHSHEVEYVLPRFEVQVVKPKMATLEE	242
Q	y		245	HLNGTITAKTYTKPVKGDVTL-----TFLPLSFWK-KKNIYTKFKINGSANFSND-	296
D	b		243	ELVATACGIYTYGKPYVGLVTKVCRKXTQSYNCHGQHSKISICEFGKQADEKGCPRQV	302
Q	y		297	-----EEMKRVMSNSLSYLDLSSRPELITTYTESVTGISKNS	339
D	b		303	VKTQVFOPROKGYDMKI EWEAKIKEDGT-----ELTGTGCEIANTLTSK-----LKFTK	353
Q	y		340	TNVFQKHDIIEFPDYTTVLK-----PSLNFATVATVRADGNQTLERRNNVITY	393
D	b		354	ANFMYRG--LFFQGVLLVDEKQPIENKRLTYQVNSVNSQFFFTDDEHGLNLTIDT	410
Q	y		394	TQNTYETVSGNSGNGKMEAVOKINYTYPOSCTKIEFPILEDSEELQKAYFLSKSS	453
D	b		411	T-NFFTSFNG-----IRVLYKONNICPDNMWVDEYHT	441
Q	y		454	MAVSHJFK--SPSEKTYOLKTRDENIKVGSPEFLV-----SGKRLKELS--YMWVR	503
D	b		442	QAHSHARITFSPEKSYIQLBELVGLTAGCQTOEIHPLLEDALIKOMKDLTFYLLIKAR	501
Q	y		504	GOLVAVGK-----ONSTFSLTPENSWTPRACIYIVYIEDDGEHISDLVKIPQL	553
D	b		502	GSIFNSGSHVFLPEQCKVGVVSPFIRVBPGMAPAKILVTLILPNEBLINDQKPIEK	561
Q	y		554	VFKOKIKLYWSKYKAPESEKVSIRISVTPDSIVGIVAVDKSVMLNANSDITMENVAHE	613
D	b		562	CFANTVNLSPFSAQSPLASDTHLTVKAT-PLSLCALTAVDQSVLLKPEAKLSPOSITYNL	620
Q	y		614	L-ELMYTGYLGMF-----NMSFVFPQEGJLAWLTD	643
D	b		621	LPOKAEAGAYLGLPYKGECENCIKADETHANGIYVTPQODLNDNDAYSVPFOSIGLFTN	660
Q	y		644	ANLTQYIDGVVD-----NAEYAEFMEBNEGHIVDIHDFSIGSS	683
D	b		681	TRVHKRPRYCPMYQAYVPLPYVGPQALMAISIRGAGY-----RSSNIRITSSMMWNGAS	733
Q	y		684	-----PHRKHPETWIMLDTMGRIRIQEVEYVPPDSITSVATGVVISEDJLG	734
D	b		734	EVAQVEVRETKYKPEPTWMDMWPDLISGDGEPLVYKVPDITIEWKASACISGTTGLG	793
Q	y		735	LTTTPELQAFQFFFLFNLPSYVSIJGEFALEITIPYVLTQKATEVUKIIEKSRPILM	794
D	b		794	LSST-ISHKVQFQFFLETLPLPSYVSGEAFILKATVYVLMHPCIRIHVLSLEMSDFLAVP	853
Q	y		795	TSSBINATGQOQLLVPSBDGATVLFPIRPTHGELPIYTVTA---LSP-----	839

```

Query Match      18.7%; Score 1374.5; DB 2; Length 1464;
Best Local Similarity 27.9%; Pred. No. 1.1e-69;
Matches 430; Conservative 243; Mismatches 56; Indels 271; Gaps 499

Qy 13 LCVCATALAAVAGPRFLVTAPGILIRPGGNVITIGVELLECPQYTVAAELIKLTASNLTVS 72
      | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 16 LALLMTAAAVHKKPQYMLVP-----TQLTVDEPKSCMLHLYIKETVYVSAS 63
      | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |

Qy 73 VLEA-----ECGVEKGSFKTLPLPLPSLPSADEIYELRYVTGTODEILFNS 119
      | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |

```

64 LMSMGKSLFSDPEVEHDLFOCVSF---TLPRSSPNEVAFLSVOIKGRTH---TFSEE 117  
Qy 120 TRLSFETKRLSVFOTDKALYKPKQOEKFRIVTLFSPFKP-YKTSILNLIKDPKSNLIQ 178  
Db 118 MAVVVRNTESTILIYQTDKPMKPGQYKFRVSVDRMLRQHHELPLVYIEDPRKNIMQ 177  
Qy 179 WLSQOSDLGVISKTFPQLSSHPILCDWSIQV-QVNDQTYOSFOVSEYLLPKFEVTLQTP 237  
Db 178 WRDITSENGIKQSLSEBPIQGFYKIMVLKSGEKEHEFSAEBEYLLPRFEQVQCPK 237  
Qy 238 YCSNNSKHLNCTINAKTYGKPVKGDVTLTFLPLSPFGKKKNIKTINKINSANFSN-- 295  
Db 238 AVTLDEKRVNTVCGCKTYGKPVLGHV-----KNTICHLFKSYFGSGYGNHYG 285  
Qy 296 ---DEEMKNVMSDNGLSSEYL-----DLSSPGVEILTYTESVTGIS----- 335  
Db 286 QGDCEBELSQOLNOCCTQAVRTSRDPMNMMNHINLOYNALITBEGTLEFYGTT 345  
Qy 336 --RNVSTNVFQKHDYIIEFPDYTVLKPSLNTATYKVTBAGNOULTEBRNNVITV 393  
Db 346 EVRKPRTNLVAVSTD-----SHFRPGIPFVQFLEBDVKG---VPVADKHIFLKV 392  
Qy 394 TORRYTEYSGSNGNOKMEAVOKINTVPOSGTFKI-----BFLIEDSSELQ 443  
Db 393 FTTKYN--SAYTDEHGLVEFSTNTTDLVAPSLTFKYKYEKGHCFFKHYCIEELKQVDF 450  
Qy 444 KAYFLGKSMVAHSLFKSPKTYIOLKTRDENIKVGSPEFLVVS---GNKRLKELS-- 497  
Db 451 VFY-----SVF-SYKSFVYLEPVTGALPCGQMHVHNYHLNGLKVLGELRBMV 498  
Qy 498 --YVWSRGOLVAVGKONSTW-----FSLTPNSWTPKACVIVYIIEDEGEIISD 545  
Db 499 FYULIMARGIIVQGTNAFMEPGLKGFNLISIPBSYMAPOMLITAILPGEVYAD 558  
Qy 546 VLKIPVQLVFNKIKLYKSKYKABSEKVSRLISVQPDISVIGIYAVDKSNLMAANDI 605  
Db 559 SAKFEIENCLLSOVGLSFRASQSLPASQTHLRVT-APPQSLCALRAVQSVLIMRPEAEL 617  
Qy 606 TMENVVHELEL-----YNTGYLGMFMN--SPAVEROEGCLW 639  
Db 618 SPSSITVNLISIKDJDSSDLNMLANFPENCKDINYDKIYSSILHRNDKOTFKFIESMGLN 677  
Qy 640 VLT-----DANLTROYIDGVYDNA--EYAEFMEBNEGHIVDIHPSLGSSPHVRKHP 691  
Db 678 IFTNLKIRDPNLCPD--DLFPQAVAIKTYD-----HTVPABVDEKPLSETVRYTYP 727  
Qy 692 ETWIMLDTNMGRYIYQEFVTVPDSITSWATGVVISSEDLGLITTPVELQAFQPPRIF 751  
Db 728 ETWIMDLVEVNVSSGVAEVAVTVPDITTEMKAGALCLPSDTGLGLSPT-ASLRAPQPFVE 786  
Qy 752 LNLPSYVIRGEBEFLAITFNYLKDATKVKYIIEKSDKFDILMTSSEINATGHQOTLLVP 811  
Db 787 LTMPSYVVRGEVFTFKATVNLNYSKCIQVSLQLEASPAFTAVPLAKD-----QDSYCL 839  
Qy 812 SEDG-ATVLPFRPTHLEIPITVTALSPAS-----DAVTOMILVKA 854  
Db 840 CEDGRQYVSMVLTPKALGNVNFVSAETQOQSELOCTEVAAPPEARKKDVTVKPLVEPE 899  
Qy 855 GIEKSYOSIILDLTNRLOSTLKLSPFPNTVTSERVOITAGDVLGSPSINGLASL 914  
Db 900 GIKKEXYMSNFLCASDAVIS--EKLSTKLPPKLVTD SARAFPSVFGDILSSSIKNTONL 956  
Qy 915 IRMPYSGEOMNINPAPNIYILDYITKKOLTDNLKEKALSPMRQGYORELLYQREDSF 974  
Db 957 LQMPYGCGBQNMVLPAPNIYVLKYNLNETDQLOETKSALIGYLSAGYORQULNYPGDSY 1016  
Qy 975 SAFG--NYDPGSGTWLSAFVLRCEFLADPYIDIDONVLAHTYTLKGHQSNGEFPDGR 1032  
Db 1017 SAFGHQNRNGNCTMTLTAFLVKTLAGARAFIFIDETHI SHAFTWLSQOKNDNGCFRSSGT 1076  
Qy 1033 VHSLELOGKNSPVTLNAYVTSLL--GYRKYQPNID-----VOESIHFLSESPSGISD 1085  
Db 1077 LFHNDLKGVDDEVTLASAYITIALLEMBLPAHPVAVSKALTCLINELE-MAKEGSNG-SY 1134

Qy 1086 NYTLALITYALSSVGS-PRAXEALNMLTWRAEGQMOPWVSSEKLSDS---WQPR--S 1139  
Db 1135 VYTKALLAAVAFALAGNODKRKEILKSLDEBAVXEDNSIHWERPOKRRPREALLYQFQAPS 1194  
Qy 1140 LDIEVAAYLALSHFLOPQTSIGIP-----IKRWLSRQNSISGFPASTODTTVALK 1189  
Db 1195 AEVEMTSYVLAHL-----TAQGTPTPEEMTSAMRIVNWITKQONSYGFSSTODTVALH 1250  
Qy 1190 ALSEPAL-----MTERNTIQTVTGPSSPPLAVQ-----PMAVNI 1228  
Db 1251 ALSRYGAATGRGTALVKGSSGTFSTKFOV-----DNSNLLLOQVSLPHIPEEYTI 1305  
Qy 1229 SANGFPAICQULNVVNVKASGSSRRRSTIQNGEAPLDV-AVXENKDDLN-HVDLNVCT 1286  
Db 1306 SVSDEGCVTYQTALKTNVFLE-----KEKYAFALQVHTVPQTCDDPKAKHRFQISL 1356  
Qy 1287 SFGGPR---SGMALMEVNLISGFMYPSEALISL---SEYKAYEYHGLNLYLDSVNET 1340  
Db 1357 EVSYTGRPASNMALVNVKMSIGFSPKSTVKMLERSDHVSRTENVKNNVLIVVDQVNE 1416  
Qy 1341 QFCVNI PAVNRFKVSNTODASVSIVDYEPBRQAVRSYNS 1380  
Db 1417 TLNFSFDVLODVEPRLKPAFVAVDYIETGEAPAFATSA 1456

Search completed: January 15, 2004, 18:14:44  
Job time : 39 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 15, 2004, 18:03:24 ; Search time 18 Seconds

(without alignments)  
3730.785 Million cell updates/sec

Title: US-10-020-095-4

Perfect score: 7348

Sequence: 1 MGGPPLTAHLICVCTAAL.....HSSVIFCFCKLYEMELML 1428

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-Processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_41.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1455	19.8	1477	1 A113_RAT	P14046 ratnus norv
2	1440	19.6	1474	1 A2MG_HUMAN	P01023 homo sapien
3	1412.5	19.2	1472	1 A2MG_RAT	P06238 ratnus norv
4	1406.5	19.1	1476	1 A2M1_MOUSE	P28665 mus musculu
5	1402	19.1	1451	1 A2M2_MOUSE	P28666 mus musculu
6	1396.5	19.0	1482	1 P2P_HUMAN	P20742 homo sapien
7	1395	19.0	1473	1 OVOS_CHICK	P20740 gallus gall
8	1373.5	18.7	1495	1 A2MG_MOUSE	P01028 mus musculu
9	1026.5	14.0	1673	1 CO3_LAMJA	P00685 lamptetra ja
10	991.5	13.5	1744	1 CO4_HUMAN	P01028 homo sapien
11	987	13.4	1738	1 CO4_MOUSE	P01029 mus musculu
12	984.5	13.4	1640	1 CO3_ONCMY	P98093 oncorhynch
13	973.5	13.2	1651	1 CO3_MAMNA	P01033 oncorhynch
14	966	13.1	1663	1 CO3_RAT	P01026 ratnus norv
15	956	13.0	1663	1 CO3_MOUSE	P01027 mus musculu
16	950	12.9	1737	1 CO4_RAT	P08649 ratnus norv
17	934	12.7	1663	1 CO3_HUMAN	P01024 homo sapien
18	921	12.5	1676	1 CO5_HUMAN	P01031 homo sapien
19	905	12.3	1680	1 CO5_MOUSE	P06884 mus musculu
20	904	12.3	1660	1 CO3_CAVBU	P12387 cavia porce
21	857.5	11.7	1620	1 CO3_EPTBU	P98094 eptaretrus
22	543	7.4	920	1 CO4_BOVIN	P01030 bos tauru
23	467.5	6.4	1536	1 Y984_THEME	P01030 bos tauru
24	432.5	5.9	726	1 CO3_FABIT	P12347 oryctolagus
25	290.5	4.0	1892	1 Y835_RICCN	P09286 rickettsia
26	283	3.9	1653	1 YFHM_ECOLI	P08408 anabaena sp
27	263.5	3.6	1906	1 YFAD_ANASP	P08408 anabaena sp
28	227	3.1	1582	1 YU30_RALSO	P08408 anabaena sp
29	187	2.5	1516	1 Y819_PSEAE	P08408 anabaena sp
30	183	2.5	1534	1 YFAS_ECO57	P08408 anabaena sp
31	180.5	2.5	1720	1 FTSH_CHUVU	P08408 anabaena sp
32	176	2.4	1534	1 YFAS_ECOLI	P08408 anabaena sp
33	174.5	2.4	2366	1 TOXB_CLODI	P18177 clostridium

34	173.5	2.4	1905	1 Y659_PASMU	Q9cmz1 paeteurella
35	166.5	2.3	2358	1 YEBJ_ECOLI	P76347 escherichia
36	164.5	2.2	2004	1 YP73_YERPE	Q8zdj2 yersinia pe
37	161	2.2	1337	1 P152_YEAST	P39685 saccharomyc
38	161	2.2	1562	1 YMA1_YEAST	Q04781 saccharomyc
39	161	2.2	2660	1 YEBJ_ECO57	Q8xv7 escherichia
40	161	2.2	3354	1 CADN_MOUSE	Q99p14 mus musculu
41	161	2.2	4349	1 FAT2_HUMAN	Q9nyq8 homo sapien
42	159.5	2.2	1609	1 FIG2_YEAST	P25653 saccharomyc
43	159	2.2	1656	1 ATC8_YEAST	Q12674 saccharomyc
44	159	2.2	2201	1 TENA_HUMAN	P24821 homo sapien
45	159	2.2	3174	1 CHAC_HUMAN	Q96f17 homo sapien

## ALIGNMENTS

RESULT 1	ID	Accession	Standard	PRT	AA
AC	P14046				
DT	01-APR-1990 (Rel. 14, Last sequence update)				
DT	15-SEP-2003 (Rel. 42, Last annotation update)				
DE	Alpha-1-inhibitor III precursor.				
OS	Rattus norvegicus (Rat).				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.				
OX	NCBI_TaxID=10116;				
RN	[1]				
RP	SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.				
RC	TISSUE=Liver;				
RX	MEDLINE=88153707; PubMed=2831216;				
RA	Brackley T.A., Northmann W., Hudson G.O., Shiels B.R., Gehring M.R.,				
RT	Rey G.H.;				
RT	"Sequence and acute phase regulation of rat alpha 1-inhibitor III				
RT	messenger RNA."				
RL	J. Biol. Chem. 263:3999-4012(1988).				
CC	- FUNCTION: PROTEASE INHIBITOR WITH A WIDE SPECTRUM OF PROTEIN				
CC	TARGETS, WHICH ATTACHES THROUGH ITS THIOLESTER FUNCTION.				
CC	- SUBUNIT: Monomer.				
CC	- MISCELLANEOUS: THIS INHIBITOR HAS REDUCED PLASMA CONCENTRATIONS IN				
CC	ACUTE PHASE.				
CC	- SIMILARITY: TO OTHER PROTEINS OF THE ALPHA-MACROGLOBULIN FAMILY,				
CC	INCLUDING COMPLEMENT COMPONENTS C3, C4, AND C5.				
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration				
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -				
CC	the European Bioinformatics Institute. There are no restrictions on its				
CC	use by non-profit institutions as long as its content is in no way				
CC	modified and this statement is not removed. Usage by and for commercial				
CC	entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a>				
CC	or send an email to <a href="mailto:license@sib-sib.ch">license@sib-sib.ch</a> ).				
CC	-----				
DR	EMBL; J03552; AAA40628.1; -				
DR	EMBL; M22358; AAA40629.1; -				
DR	PIR; A29952; A29952.				
DR	HSP; P01023; 1BV8.				
DR	GO; GO:001714; F:wide-spectrum protease inhibitor activity; NAS.				
DR	InterPro; IPR002890; A2M_N.				
DR	InterPro; IPR001599; MacroglloblnA2.				
DR	Pfam; PF00207; A2M; 1.				
DR	Pfam; PF01835; A2M_N; 1.				
DR	PROSITE; PS00477; ALPHA_2_MACROGLOBULIN; 1.				
KW	Serine protease inhibitor; Glycoprotein; Plasma; Bait region; Signal;				
KW	Liver; Thioester bond.				
FT	SIGNAL	1	24		
FT	CHAIN	25	1477		ALPHA-1-INHIBITOR III.
FT	DOMAIN	601	750		BAIT REGION (APPROXIMATELY).
FT	DISULFID	48	86		BY SIMILARITY.
FT	DISULFID	251	295		BY SIMILARITY.
FT	DISULFID	269	283		BY SIMILARITY.
FT	DISULFID	468	563		BY SIMILARITY.

FT DISULFID 595 774 BY SIMILARITY.  
 FT DISULFID 643 678 BY SIMILARITY.  
 FT DISULFID 850 886 BY SIMILARITY.  
 FT DISULFID 924 1324 BY SIMILARITY.  
 FT DISULFID 1082 1130 BY SIMILARITY.  
 FT DISULFID 1355 1470 BY SIMILARITY.  
 FT CROSSLINK 975 978 IsoleutamyI cysteine thioester (Cys-gln).  
 FT CARBOHYD 55 55 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 247 247 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 301 301 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 321 321 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 393 393 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 508 508 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 750 750 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 777 777 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 872 872 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 994 994 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 1143 1143 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 1314 1314 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 1427 1427 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT VARIANT 4 4 D -> N.  
 FT VARIANT 87 87 V -> L.  
 FT VARIANT 137 137 M -> V.  
 FT SEQUENCE 1477 AA; 163773 MW; 4DC05367C8385D2B CRC64;

Query Match 19.8%; Score 1455; DB 1; Length 1477;  
 Best Local Similarity 29.4%; Pred. No. 8,4e-72;  
 Matches 457; Conservative 237; Mismatches 578; Indels 284; Gaps 50;

Qy 13 LCVCTAALVAPGPRFLVAPGILRPGN---VTIGVELLEHCPQOVTAKEALLKTAASN 68  
 Db 9 LCLSSALLAPLPFASL-----NNGSKYMLVPSQLYTERPEKICLHLVHLNETVT 59  
 Qy 69 LTVVLEAG-----VFEGSKFK--TLTLPILPLNSADEIYELARTGTOBELLEPN 118  
 Db 60 VTALLISQGRGKLFDELIVADKLFHCVSFTIPRLPSSSEESLDINIEGAKHK--FSE 116  
 Qy 119 STRLSEFKRISVFQTDKALYKPKQEVKFRIVTLFSPFKYKTSINLT-----IKPKS 173  
 Db 117 RRVLVKKKESVVFQTDKPKMYKQSVKFRVNSMDKRLHP-----LNEFLPLAIIEDPKM 172  
 Qy 174 NLIQMLSQSGDLGVISTKTPQLSHPILGDMSI-----QVQVNDQTYQSFQVSEVYLKPF 229  
 Db 173 NRIQMOMQVTENGLKQLSFSLSEPIQGYKIVLKQSVKE---HSFTVMEFVLPRF 229  
 Qy 230 EVTLQTPLYCSMNSGHLNGITIAKTYGPKYKGVTLTFLPLSTFWGKKGNITKFKINGS 289  
 Db 230 GVDKVPVAISVDEIINVTACATYTGKVPVGHVKISL-----CHGN 272  
 Qy 290 ANFSFNDEMGVNDMS---SNGLS-----EYLDLSSPGVEILTTVESVYG 333  
 Db 273 PTFSEETKSGCKEDSRLDNNGCSTQEVNITEFQLEKENYILQMHC--AFHVAIVTTEBGTG 330  
 Qy 334 ISRVNSTVNFVKQDHYIEFPDYTTVLKPSLNFATYAVKTRADGNQUTLEERRNNVITV 393  
 Db 331 SEFGSGRIEYERFRNKPLFLKADSHFRGIPFVKVYLVDIKDPIDPNEQ---VLTKA 386  
 Qy 394 TORVYTEWSSGNSGNQMEAVOKINTVPOSSTFKIEFLPLESSELQKAYFLGSKSS 453  
 Db 387 RDAGYTAATTDHGLAK-----FSIDTNGISDYS--INIKVYH--KEES. 427  
 Qy 454 MAVHS-----LFKSPSKTYIQLKTR-----DENIKVGSPEEL--VVSQN 490  
 Db 428 SCHSSCTAEHAEAHHTAIVAVYSLSKSYITLDBAGVLPENQIHTVOAHFELGQVLGV 487  
 Qy 491 KRLKELSYMVVSRQGLVAVGK-----QNSTFSLTPENKSTPKACVIVYIYED 538  
 Db 488 LQGVLFHVLVWAQSSIIQGTGNHTQOVERGESQVGNFLERLPEFMSVPAVKMLIYILP 547  
 Qy 539 DGEIISDVLKLPVQVLFQNKIKLYSKVKAEPSEKVSIRISVTPDLSIGVIAVDKSVNL 598  
 Db 548 DGEVIAASVVKQVKEKLCRNKVLHLSFSPQSLPASQTHMRVTAS--POSICGLRAVDQSVTL 606

Qy 599 -----MNASNDITMENNVVHELELYNTGY--LGMFNNSFAVFOECGLMV 640  
 Db 607 QKPEAELEPSLIYDLPEMQDSNFIASNDPDEDYCLMQPIAREKDYRVYRETGLMA 666  
 Qy 641 LTTDANL-TKQYINGVYDNAEYA-----ERFMEEGHLYVDIHDFLSSP----- 684  
 Db 667 FTMKIKLPTCYCNDYDMVLAVPAVALDSTGRGWESL--PVAVKGS--PLPQEPKRPDP 724  
 Qy 685 -----HYRKHPETMILDTNMGYRIYQEPREVTVPDSITSMVATGFISEDLGLT 736  
 Db 725 PPKDPVETIRNTPPETIMWDLTVANSSGVTELEMTVPDITTEKACALCLSDTGLGLS 784  
 Qy 737 TTPVELQAPQPFILFLNLPYSVIRGEFALETIFNYLKDATEVKYIIEKSKFDILMTS 796  
 Db 785 SV-ASFAQPFPEVLELMPYSVIRGEAFTLKATVNLVLPFLMAVVLLEASPDF----- 837  
 Qy 797 SEINATGHQOTLVPSFDATVLPPIRPTHLGELPIITVTLSS-----PTAS----- 842  
 Db 838 TAVPVENNQDSYCLGANGRHTSSMLVTPKSLGNVNSVSAAEQSPQPCGEVATVPETG 897  
 Qy 843 --DAVTQMLVKAEGIEKYSOSILDLTDNRLOSTLKTLSFSPPTVTYGSERVOITAI 900  
 Db 898 RKDTVAKVILVEBGLKKEHTFSSLLCASDAEIS--ETLSLLPLPTVVKOSARAHFSVM 954  
 Qy 901 GDVLGSPINGIASLIRMPYGCGEQNMIFAPNTIYLDLYTKKQUTDNLKEKALSPNRQ 960  
 Db 955 GDILSSAKNTQNLQMPYGCGEQNMVLFAPNTIYLYKLTNETQQLTEKISKALGYLRAG 1014  
 Qy 961 YQRELLYQREDSFSAFGNYD--PSGSTWLSAVYLCFLEADYIDIDQVNLHRTYWLK 1018  
 Db 1015 YQRELLYQREDSFSAFGNYD--PSGSTWLSAVYLCFLEADYIDIDQVNLHRTYWLK 1074  
 Qy 1019 GHQKSGEPMDFPRVHISELGNGKSPVTLTAYIVTSLIGRYKQPNID--VOESIHFFE 1076  
 Db 1075 KQKDSCEFTSSGSLNNAKGGVDEITISATITMAL--ESSLPDTDVVSKALSCLE 1132  
 Qy 1077 SEFSR-----GISDNTYLAITYALSSVGS--PRAKALNMLTVRABQEGMQPW----- 1124  
 Db 1133 SSMENIEQGGNGSFVYTKKALMAVAPLAGQEKRENLKSLDKEXALKEDNSIMHERQKP 1192  
 Qy 1125 VSSBSKLSDSWQPRSLDIEVAAYALLSHLQFOTSGIP-----IMKULSROKNS 1174  
 Db 1193 TKSEGYLYTP-QASSAEVEMSAVYLAFL--TAQAPAPEDIALSMGTIKMLTKQONS 1247  
 Qy 1175 LGGPASTQDPTVLLKLSER--AALMTERTNIOVTYTGSPSPPLAVVQ----- 1222  
 Db 1248 YGGFSSQDPTVVALDLSTKGAATFSKQKTPSVTYQSSGFSQKQOVKSNRLLOQVS 1307  
 Qy 1223 ---PMAVNISANGFGFALCOLNVVNVKASGSSRRRSIQONQAPDLVAVK---ENK 1274  
 Db 1308 LPYIPGNYTVSGBGCVVAQTLRLRNVPLE-----KQPAFALKVQTVPLTCNNP 1358  
 Qy 1275 DDINHDVAVCTSPSGP--GRSGMALMEVNLISGFMPVSEALISSEYKYK-----E 1324  
 Db 1359 KQONSFOISLEISYMGSRPASNNVIADVKKLSGF-----IPLKPTVKLERLGHVSRTE 1412  
 Qy 1325 YDHGKNTLYDSVNEHQFCVNIYAVNRFKYSNQDASVSLVDYEPBRQAVRSYNS 1380  
 Db 1413 VTTNNVLLVLDQVNTQTLSPFSFIQODIPVKNQPAIVKVDYETDEVAFAEYSS 1468

RESULT 2  
 A2MG\_HUMAN  
 ID A2MG\_HUMAN STANDARD; PRT; 1474 AA.  
 AC P01023; Q13677;  
 DT 21-JUL-1986 (Rel. 01, Created)  
 DT 13-AUG-1987 (Rel. 05, Last sequence update)  
 DT 15-SEP-2003 (Rel. 42, Last annotation update)  
 DE Alpha-2-macroglobulin precursor (Alpha-2-M).  
 GN A2M.  
 OS Homo sapiens (human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

OK NCBI\_TaxID=9606;  
RN (1)  
RP SEQUENCE FROM N.A.  
RX MEDLINE=85190481; PubMed=2581245;  
RA Kan C.-C., Solomon E., Belt K.T., Chain A.C., Hiorne L.R., Fey G.H.;  
RT "Nucleotide sequence of cDNA encoding human alpha 2-macroglobulin and  
RT assignment of the chromosomal locus."  
RT Proc. Natl. Acad. Sci. U.S.A. 82:2282-2286(1985).  
RN (2)  
RP SEQUENCE FROM N.A., AND VARIANT HIS-704.  
RX MEDLINE=92246939; PubMed=1374237;  
RA Mathijs G., Devriendt K., Cassiman J.-J., van den Berghe H.,  
RA Marynen P.;  
RT "Structure of the human alpha-2 macroglobulin gene and its promoter."  
RT Biochem. Biophys. Res. Commun. 184:596-603(1992).  
RN (3)  
RP SEQUENCE OF 24-1474, AND CHARACTERIZATION.  
RX MEDLINE=84239807; PubMed=6203908;  
RA Sottrup-Jensen L., Stepanik T.M., Kristensen T., Wierzbicki D.M.,  
RA Jones C.M., Loenblad P.B., Magnusson S., Petersen T.E.;  
RT "Primary structure of human alpha 2-macroglobulin. V. The complete  
RT structure."  
RT J. Biol. Chem. 259:8318-8327(1984).  
RN (4)  
RP ERRATUM.  
RA Sottrup-Jensen L., Stepanik T.M., Kristensen T., Wierzbicki D.M.,  
RA Jones C.M., Loenblad P.B., Magnusson S., Petersen T.E.;  
RT J. Biol. Chem. 260:6500-6500(1985).  
RN (5)  
RP SEQUENCE OF 832-1474 FROM N.A.  
RC TISSUE=Liver;  
RX MEDLINE=85219061; PubMed=2408344;  
RA Bell G.I., Ball L.B., Sanchez-Pescador R., Merryweather J.P.,  
RA Scott J., Eddy R.L., Shows T.B.;  
RT "Human alpha 2-macroglobulin gene is located on chromosome 12."  
RT Somat. Cell Mol. Genet. 11:285-289(1985).  
RN (6)  
RP SEQUENCE OF 672-747.  
RX MEDLINE=90242963; PubMed=1692292;  
RA Marynen P., Devriendt K., van den Berghe H., Cassiman J.J.;  
RT "A genetic polymorphism in a functional domain of human pregnancy  
RT zone protein: the bait region. Genomic structure of the bait domains  
RT of human pregnancy zone protein and alpha 2 macroglobulin."  
RT FEBS Lett. 262:349-352(1990).  
RN (7)  
RP INHIBITORY SITE.  
RX MEDLINE=84030513; PubMed=6195065;  
RA Virca G.D., Salvesen G.S., Travis J.;  
RT "Human neutrophil elastase and cathepsin G cleavage sites in the bait  
RT region of alpha 2-macroglobulin. Proposed structural limits of the  
RT bait region."  
RT Hoppe-Seyler's Z. Physiol. Chem. 364:1297-1302(1983).  
RN (8)  
RP INHIBITORY SITE.  
RX MEDLINE=81212827; PubMed=6165619;  
RA Sottrup-Jensen L., Loenblad P.B., Stepanik T.M., Petersen T.E.,  
RA Magnusson S., Joensuu H.;  
RT "Primary structure of the 'bait' region for proteinases in alpha 2-  
RT macroglobulin. Nature of the complex."  
RT FEBS Lett. 127:167-173(1981).  
RN (9)  
RP INHIBITORY SITE.  
RX MEDLINE=81255805; PubMed=6167263;  
RA Hall P.K., Nelles L.P., Travis J., Roberts R.C.;  
RT "Proteolytic cleavage sites on alpha 2-macroglobulin resulting in  
RT proteinase binding are different for trypsin and Staphylococcus  
RT aureus V-8 proteinase."  
RT Biochem. Biophys. Res. Commun. 100:8-16(1981).  
RN (10)  
RP INHIBITORY SITE.  
RX MEDLINE=82095610; PubMed=6172288;  
RA Mortensen S.B., Sottrup-Jensen L., Hansen H.F., Petersen T.E.,  
RA Magnusson S.;

RT "Primary and secondary cleavage sites in the bait region of alpha 2-  
RT macroglobulin."  
RT FEBS Lett. 135:295-300(1981).  
RN (11)  
RP STRUCTURE BY NMR OF 1337-1474.  
RX MEDLINE=99081553; PubMed=9865955;  
RA Huang W., Dolmer K., Liao X., Gettins P.G.W.;  
RT "Localization of basic residues required for receptor binding to the  
RT single alpha-helix of the receptor binding domain of human alpha2-  
RT macroglobulin."  
RT Protein Sci. 7:2602-2612(1998).  
RN (12)  
RP VARIANT ILE-1000.  
RX MEDLINE=91187639; PubMed=1707161;  
RA Poller W., Faber J.-P., Olek K.;  
RT "Sequence polymorphism in the human alpha2-macroglobulin (A2M) gene."  
RT Nucleic Acids Res. 19:198-198(1991).  
RN (13)  
RP VARIANTS TYR-972 AND ILE-1000.  
RX MEDLINE=92128897; PubMed=1370808;  
RA Poller W., Faber J.-P., Klobbeck G., Olek K.;  
RT "Cloning of the human alpha 2-macroglobulin gene and detection of  
RT mutations in two functional domains: the bait region and the  
RT thiolester site."  
RT Hum. Genet. 88:313-319(1992).  
RN (14)  
RP FUNCTION: IS ABLE TO INHIBIT ALL FOUR CLASSES OF PROTEINASES BY A  
RN UNIQUE "TRAPPING" MECHANISM. THIS PROTEIN HAS A PEPTIDE STRETCH,  
RN CALLED THE "BAIT REGION" WHICH CONTAINS SPECIFIC CLEAVAGE SITES  
RN FOR DIFFERENT PROTEINASES. WHEN A PROTEINASE CLEAVES THE BAIT  
RN REGION, A CONFORMATIONAL CHANGE IS INDUCED IN THE PROTEIN WHICH  
RN TRAPS THE PROTEINASE. THE ENTRAPPED ENZYME REMAINS ACTIVE AGAINST  
RN LOW MOLECULAR WEIGHT SUBSTRATES (ACTIVITY AGAINST HIGH MOLECULAR  
RN WEIGHT SUBSTRATES IS GREATLY REDUCED). FOLLOWING CLEAVAGE IN THE  
RN BAIT REGION A THIOLESTER BOND IS HYDROLYZED AND MEDIATES THE  
RN COVALENT BINDING OF THE PROTEIN TO THE PROTEINASE.  
RN (15)  
RP SUBUNIT: HOMOTETRAMER, WHICH CONSISTS OF TWO PAIRS OF DISULFIDE-  
RN LINKED CHAINS.  
RN (16)  
RP TISSUE SPECIFICITY: Plasma.  
RN (17)  
RP MISCELLANEOUS: CONTRARY TO THE RAT PROTEIN, WHICH IS AN ACUTE  
RN PHASE PROTEIN, THIS PROTEIN IS ALWAYS PRESENT AT HIGH LEVELS IN  
RN CIRCULATION.  
RN (18)  
RP SIMILARITY: TO OTHER PROTEINS OF THE ALPHA-MACROGLOBULIN FAMILY,  
RN INCLUDING COMPLEMENT COMPONENTS C3, C4, AND C5.  
RN (19)  
RP This SWISS-PROT entry is copyright. It is produced through a collaboration  
RN between the Swiss Institute of Bioinformatics and the EMBL outstation -  
RN the European Bioinformatics Institute. There are no restrictions on its  
RN use by non-profit institutions as long as its content is in no way  
RN modified and this statement is not removed. Usage by and for commercial  
RN entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
RN or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
RN (20)  
RP EMBL, M11313; AAA1551.1; -  
RN EMBL, Z11711; CAA77774.1; -  
RN EMBL, M36501; AAA51552.1; -  
RN EMBL, X68728; CAA48670.1; -  
RN EMBL, X68729; CAA48670.1; JOINED.  
RN PIR; A94033; MAHU.  
RN PDB; 1BV8; 30-SEP-98.  
RN SWISS-2DPAGE; P01023; HUMAN.  
RN Genew; HGNC:7; A2M.  
RN MIM; 103950;  
RN GO; GO:0016975; F:alpha-2 macroglobulin; NMS.  
RN InterPro; IPR002890; A2M\_N.  
RN InterPro; IPR001599; Macrogloblna2.  
RN Pfam; PF00207; A2M; 1.  
RN Pfam; PF01835; A2M\_N; 1.  
RN PROSITE; PS00477; ALPHA\_2\_MACROGLOBULIN; 1.  
RN Serine protease inhibitor; Glycoprotein; Plasma; Bait region; Signal;  
RN 3D-structure; Polymorphism; Thiolester bond.  
RN SIGNAL 1 23  
RN CHAIN 24 1474 ALPHA-2-MACROGLOBULIN.  
RN FT 693 694 TRANSGUTAMINASE CROSS-LINKING SITE

```

FT  DOMAIN      690      728      (POTENTIAL)
FT  SITE        704      709      BAIT REGION.
FT  SITE        719      723      INHIBITORY.
FT  SITE        730      735      INHIBITORY.
FT  DISULFID    48       86
FT  DISULFID    251     299
FT  DISULFID    269     287
FT  DISULFID    278     431
FT  DISULFID    470     430
FT  DISULFID    563     563
FT  DISULFID    595     771
FT  DISULFID    642     689
FT  DISULFID    821     849
FT  DISULFID    847     883
FT  DISULFID    921     1321
FT  DISULFID    1079    1127
FT  DISULFID    1352    1467
FT  CROSSLINK    972     975
FT  CARBOHYD     55      55
FT  CARBOHYD     70      70
FT  CARBOHYD    247     247
FT  CARBOHYD    396     396
FT  CARBOHYD    410     410
FT  CARBOHYD    869     869
FT  CARBOHYD    991     991
FT  CARBOHYD   1424     1424
FT  VARIANT      704      704
FT  VARIANT      972      972
FT  VARIANT     1000     1000
FT  CONFLICT     63      63
FT  CONFLICT    563     563
FT  CONFLICT    1148    1148
FT  CONFLICT    1195    1195
FT  STRAND      1341    1347
FT  TURN        1352    1353
FT  HELIX       1355    1359
FT  STRAND      1360    1369
FT  STRAND      1379    1384
FT  STRAND      1390    1391

Query Match      19.6% Score 1440; DB 1; Length 1474;
Best Local Similarity 29.1%; Pred. No. 5.6e-71;
Matches 451; Conservative 266; Mismatches 551; Indels 284; Gaps 59;

Isoglutamy] cysteine thioester (Cys-Gln).
N-LINKED (GLCNAC. . .)
N-LINKED (GLCNAC. . .)
N-LINKED (GLCNAC. . .)
N-LINKED (GLCNAC. . .)
N-LINKED (GLCNAC. . .)
N-LINKED (GLCNAC. . .)
N-LINKED (GLCNAC. . .)
N-LINKED (GLCNAC. . .)
N-LINKED (GLCNAC. . .)
N-LINKED (GLCNAC. . .)
R -> H (IN dbSNP:1800434).
/FTid=VAR_000012.
C -> Y (PROBABLY INTERFERES WITH THE
ACTIVITY; dbSNP:1800433).
/FTid=VAR_000013.
V -> I (IN dbSNP:669).
/FTid=VAR_000014.
MISSING (IN REF. 3).
C -> E (IN REF. 3).
A -> D (IN REF. 5).
H -> D (IN REF. 5).

```

```

Qy  332 TGISRVSTNVFVKQHDYIIFFPDYTVLKPSTLNTATYKATRADGNQLTLEBRNNVVI 391
Db  348 TKLS-FVKVDSHFRQG---IPFFG-----QVRLVDEKGVPIP---NKVI 384
Qy  392 TVYQRYVTEYSGNSGNQMGVAO-KINYT--VPQSGTFKIEFF-----ILED 438
Db  385 FT-RGNEANY--SNATTBHGLVQPSINTTNVGTSLTVRVVYKORSPCYQWVSEEH 441
Qy  439 SELQKAYFLGSKSMAVHSLFKSPKTYIQLKTRDENIKVGS-----PEELVSGNK--R 492
Db  442 EBAHHTAY-----LVSPSKSYHLEPMSHELPCGHTQYVQAHYILNGTLLG 489
Qy  493 LKELS--YVVSQGLVANG-----KQNST---MFLTPENSWPKACVIVYIEDDG 540
Db  490 LKLSFYLLIMAKGIVRGTHGLLVKQEDMKGHFISIPVKSDIAPVARLLIYAVLPFG 549
Qy  541 EIIISDLKIPVQLVFPKQIKLYMSKTKABSEKVSRLISTQPDLSVGIYAVDKSVNLMN 600
Db  550 DVIQSAKTDVENCLANKVDLSFSPQSLPASHAHLRVTA--PQSVCALRAVDQSVLLMK 608
Qy  601 ASNDITMENVHBL-ELYNTGY-----YLGFMNNS-----FA 631
Db  609 PDLELSASSVYNLLPEKDLTFPGPLNDODDEICIRHNVYINGITYTPVSTINEKDMVS 668
Qy  632 VPQECGLMWLTDAANTKDYIDGYDNAE-----YAEFMEENEGHIVDIDFSL 680
Db  669 FLBDMGLKAFNTSKIRKPKCPQLQOYEMHGFGLRVGFESDVMDRGHARLVHVEE--- 725
Qy  681 GSSPH---VRKHPEPWIMLDTMGRIYIOEPVYTPDITSVNTATGFISELGLGLTT 737
Db  726 ---PHETVAKYPEPWIMDLVNVNSAGVAEVGVTPDITTEKKAFCISEBAGIGISS 782
Qy  738 TPVELAOPFPFLMLPYSVIRGEEFALBITFIYLIKATEVYKIIKSDKEDIMTSS 797
Db  783 T-ASLAPQFPFELTMPYSVIRGEAFTLKATYANTLPRCIRSVOLDSRAFLAVPVER 841
Qy  798 E-----INATGQOQLLVPSBDGATVLPPIRPHLGEIDITYTALS-----PTA 841
Db  842 EQAPHCICANGRO-----TVSWAVTPKSLGNVNFVSAEALSESQELCGTEVPSV 890
Qy  842 -----SDATYQMLLVAAEGIEKYSOSILLDLTNLQSTLKLTSFPFNTYTGSERV 896
Db  891 PEHGRDVTYIKPLVPEGLEKETTFNSLDCSGGSEVS---EELSLEKLPNVVEEARAS 947
Qy  897 ITAIGDVLPSINGLASLIRMPYCGEOMNINAPINYLIDYTKKQOLTDLNLEKALSP 956
Db  948 VSVLGDILSANOQNTQNLQMPYCGEONNVLPAPNIVUDYNEFQOLTPEVKSALIGY 1007
Qy  957 MRQYORELLYOREDESFSAFG-NYDPS-GSTWLSAFVLRCELEADPYIDIDQNVLAHRY 1014
Db  1008 LMTGYQRLNLYKHVYDGSYTFGERYGRNQNTWLTAFVLTFAQAAVYIFIDEAHITQAL 1067
Qy  1015 TWLKHQKSGEEMWDGRIYHSELQGNKSPVTLTYITSLIGYKIQPNIDVQESHIF 1074
Db  1068 IWTISQROKNGCFRRSGSLNNAIKGVDEVTLSAYITALELPLTVTHPVRNALFC 1127
Qy  1075 LES-----EFSKGISDNTYLLITVALSYLSVS-PAKEXALMNLTRABOEGMQW--- 1124
Db  1128 LESAMKTAQBGDHG-SHYTTKALAYAFALAGQDRKKVYLSLNEAYAKKONSVMWERP 1186
Qy  1125 VSSSEKLSDSWQPR--SLDIEVAAYALLSHPLQFQ---TSEGIP---IMRWLSROQNSL 1175
Db  1187 QKRAVGHVPEHQASAEVEMTSYVLLA-YLTAQAPAPSEDLTSAITVKNWITKQONNA 1245
Qy  1176 GGFASQDTTVALKALSEFALMNTERTNIQYTVTGSSPSPLAVVQ-----1222
Db  1246 GGFSSQDTTVVALLHALSKYGAATFT-RTGKAQVITQSSGTFSSKQVQVNNRLLQOVS 1304
Qy  1223 ----PMAVNISANGFEFALCOLNVVNVYVAKASGSSRRRRSISIQNEAFDLDVAVK---END 1275
Db  1305 LPELPEGYSMKYTGECVYLQTSLLKNTI-----LPEKEEFPFALGVQTLDPQTC 1353
Qy  1276 DLN-HVDLNVCTSFSGPG---RSGMALMEVNLSGFMVPSAISTL---SETVAKVEYDHG 1328

```

DB 1354 EPKAHSFOISUSVSTGSRSSANNAIVGVKAVSGTIPKPTVKLERNHNSRREVSSN 1413

QY 1329 KNTLYDSVNETQFCVNIPAVNFKVNTQDASVSTVDYEPBROAVRGSN 1380

DB 1414 HVLITDKKSNQTLSTFTVLQDVPVPRDKPAIVKYVDYETDERPAIAYNA 1465

RESULT 3

A2MG RAT STANDARD; PRT; 1472 AA.

AC P06238; 01-JAN-1988 (Rel. 06, Created)

DT 01-JAN-1988 (Rel. 06, Last sequence update)

DT 15-SEP-2003 (Rel. 42, Last annotation update)

DE Alpha-2-macroglobulin precursor (Alpha-2-M).

GN A2M.

OS Rattus norvegicus (Rat).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.

OX NCBI\_TaxID=10116;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=Liver;

RX MEDLINE=87083488; PubMed=2432068;

RA Gehring M.R., Shiels B.R., Northern W., de Bruijn M.H.L., Kan C.-C., Chain A.C., Noonan D.J., Fey G.H.,

RT "Sequence of rat liver alpha 2-macroglobulin and acute phase control of its messenger RNA."

RL J. Biol. Chem. 262:446-454(1987).

RN [2]

RP SEQUENCE OF 1-164 FROM N.A.

RC STRAIN=Wistar; TISSUE=Liver;

RX MEDLINE=89160243; PubMed=2466233;

RA Kunz D., Zimmermann R., Heisig M., Heinrich P.C.,

RT "Identification of the promoter sequences involved in the interleukin-6 dependent expression of the rat alpha 2-macroglobulin gene."

RL Nucleic Acids Res. 17:1121-1138(1989).

RN [3]

RP SEQUENCE OF 178-227 AND 420-526 FROM N.A.

RX MEDLINE=86033908; PubMed=2414291;

RA Hayashida K., Okubo H., Noguchi M., Yoshida H., Kangawa K., Matsuo H., Sakaki Y.;

RT "Molecular cloning of DNA complementary to rat alpha 2-macroglobulin mRNA."

RL J. Biol. Chem. 260:14224-14229(1985).

CC -1- FUNCTION: IS ABLE TO INHIBIT ALL FOUR CLASSES OF PROTEINASES BY A UNIQUE "TRAPPING" MECHANISM. THIS PROTEIN HAS A PEPTIDE STRETCH, CALLED THE "BAIT REGION" WHICH CONTAINS SPECIFIC CLEAVAGE SITES FOR DIFFERENT PROTEINASES. WHEN A PROTEINASE CLEAVES THE BAIT REGION, A CONFORMATIONAL CHANGE IS INDUCED IN THE PROTEIN WHICH TRAPS THE PROTEINASE. THE ENTRAPPED ENZYME REMAINS ACTIVE AGAINST LOW MOLECULAR WEIGHT SUBSTRATES (ACTIVITY AGAINST HIGH MOLECULAR WEIGHT SUBSTRATES IS GREATLY REDUCED). FOLLOWING CLEAVAGE IN THE BAIT REGION A THIOLESTER BOND IS HYDROLIZED AND MEDIATES THE COVALENT BINDING OF THE PROTEIN TO THE PROTEINASE.

CC -1- SUBUNIT: HOMOTETRAMER, WHICH CONSISTS OF TWO PAIRS OF DISULFIDE-LINKED CHAINS.

CC -1- TISSUE SPECIFICITY: Plasma.

CC -1- INDUCTION: BY INFLAMMATORY STIMULUS THE LEVEL OF THIS PROTEIN FIRST INCREASES, THEN DECREASES AFTER A MAXIMUM.

CC -1- SIMILARITY: TO OTHER PROTEINS OF THE ALPHA-MACROGLOBULIN FAMILY, INCLUDING COMPLEMENT COMPONENTS C3, C4, AND C5.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/) or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL; J02635; AAA40636.1; -.

DR EMBL; X13983; CA32164.1; -.

DR EMBL; X13984; CA32164.1; JOINED.

DR EMBL; X13985; CA32164.1; JOINED.

DR EMBL; M11792; AAA40637.1; -.

DR EMBL; M11793; AAA40638.1; -.

DR PIR; A26122; A26122.

DR HSSP; P01023; 1BV8.

DR GO; GO:0016975; F:alpha-2 macroglobulin; NMS.

DR InterPro; IPR002890; A2M\_N.

DR InterPro; IPR001599; Macrogloblna2.

DR Pfam; PF00207; A2M; 1.

DR Pfam; PF01835; A2M\_N; 1.

DR PROSITE; PS00477; ALPHA\_2 MACROGLOBULIN; 1.

KW Serine protease inhibitor; Glycoprotein; Plasma; Bait region; Signal; Thioester bond.

FT SIGNAL 1 27

FT CHAIN 28 1472

FT DOMAIN 620 750

FT DISULFID 52 90

FT DISULFID 254 302

FT DISULFID 272 290

FT DISULFID 281 434

FT DISULFID 473 473

FT DISULFID 598 769

FT DISULFID 647 694

FT DISULFID 819 847

FT DISULFID 845 881

FT DISULFID 919 1319

FT DISULFID 1077 1125

FT DISULFID 1350 1465

FT CROSSLINK 970 973

FT CARBOHYD 59 59

FT CARBOHYD 74 74

FT CARBOHYD 250 250

FT CARBOHYD 399 399

FT CARBOHYD 651 651

FT CARBOHYD 772 772

FT CARBOHYD 867 867

FT CARBOHYD 989 989

FT CARBOHYD 1364 1364

FT CARBOHYD 1422 1422

FT CARBOHYD 1426 1426

FT CONFLICT 23 23

FT CONFLICT 490 490

SEQ SEQUENCE 1472 AA; 163701 MW; D08FA99AEB10FE0B CRC64;

Query Match 19.2%; Score 1412.5; DB 1; Length 1472;

Best Local Similarity 27.6%; Pred. No. 1.8e-69;

Matches 426; Conservative 275; Mismatches 580; Indels 265; Gaps 51.

QY 5 PLTNAHLICVCTAALAVAPGPRPLVTPAGIIRPGAVTIGVELHEHCPSOVTVKAEILK 64

DB 13 PLTNAHLICVCTAALAVAPGPRPLVTPAGIIRPGAVTIGVELHEHCPSOVTVKAEILK 70

QY 65 TASN--LTVSVLAEAGVEKESFKTLTPSLPLNSADSI--YELRYNGRTODELFSNST 120

DB 71 VHNGSLFTDLVVDLDLPHCTSF-----TVQSSDDELMFPTVOVKATHE--FRROS 121

QY 121 RLSPETKRIISVFIQTDKALYKPKQEVKFRITVLSDFPKRYTSNLIIL-IXDKPSNLIQOM 179

DB 122 TVLVYKKESLVFAQDKRIYKPGQVVRVAVSLDESFRHMLPLVLTIOQPKNNRIQOM 181

QY 180 LSQOSDLVISTKPOLSHPLIGDMSIOVQV-NDQTYQSQVSEYVLPKREVTLTQTPLY 238

DB 182 QNFNEGKIKOLSFPLSESPFQSGYKIVIRTESGRVTHPSVEFVLPKREVTLTQTPLY 241

QY 239 CSMSKHLNGITTKATYTGKPVKGVNLTFL-----PLSFPGKKKNI----- 280

DB 242 ITLIEENNVASCGIYTGKVPKGVTVNIRKYSNPNCGEESVACEKLSQQLDGRG 301

QY 281 -----TKTFKINSANFSSFNDEEMKQVMSNGLSFYLDLSSGPFVEIITVTVESYTG 333

```

Db 302 CFSQLVKTSFOLK-ROEYEMOLDVHAKIOEGTGVEE---TGKG---LTKITRTITK 352
Qy 334 ISRVSTVTFKQHDYIEFPDYTTVLKPSLNFATATVTRADGNQGLTEERANNVITV 393
Db 353 LS-----FVNDSHFRQGIPIFGQVILVUGRGPIIP----- 383
Qy 394 TORNYTEWSSGNSGNQMEAVQKINITYVPOSGTFKLEFPILLED--SSELQKAYF-- 447
Db 384 ---YETIFIGAD-----FANLYINTTDDKGLAR--FSINTDIMGSTLTVAKYKDS 431
Qy 448 -----LGSKSMVAHSLFK--SPKTYIQLKTRDINIKVSGPFEL---VSGN--K 491
Db 432 NACGFWLTEENEAHTAYAVFSPSPRSFLHLSLPKDCDTLEQAHYIINGEAMQ 491
Qy 492 RLKEL--SYMVVSRGOLVAVG-----KONSTM--FS--LTPENSWPKACVITYIEDD 539
Db 492 ELKELVEFYLLMAAGIYRAGTHVLPKQGMGRHFSILLIMETDLAVAVLVYALPN 551
Qy 540 GEISDVLIKIPVQVFNKKIKLYWSKYAPSEKVSLSISVTPDSIVGIVAVDKSVNL 599
Db 552 GEVVDGFAKYEIENCLANKVDLVRPNNGLPATALLSV--MASPOSICGLRAVDQSVLM 610
Qy 600 NASDITMENV-----VHELELYNTGYILGEMFNS--- 629
Db 611 KPTELASLIYDLIPVNDLTGFPQADQREEDTNGCVKQNDTYINGILYSPVONTNEED 670
Qy 630 -FAVFOECGLWLTLDANLTQDYI--DGVYDNAEYARFEMEENEGHVIDHDSLSSPHVR 687
Db 671 MYGCLMGKLVFNINSIRKRCVCERLDNKGIPTAHTLVSQSHDAFLSESSPTEFRR 730
Qy 688 KHPEWTIMDLTNGYRIYOEFEVTPDSITSWATGFEVISEDGLGLTTPVPELQAFOP 747
Db 731 SYPEWTIMDLVVDVDSAGVAEVEVTPDTITTEWAKGACLSNDTGLGLSPV--VQFOAFOP 789
Qy 748 PFILNLPYSIRREPELETTIRNYLKDAVEYVILKSKPKPILMTSSINATGHOOT 807
Db 790 FEVELTMEYSVIRBAPFLKATVNTYLPCTIRVAVOLEASP--DELAPEKEORSH-- 844
Qy 808 LLVSEDAVLPFIRPFLHGEIPITVTALSPTA-----SDAVTOMILV 851
Db 845 -CICMNRHTASMAVIPKSLGNVFTVSALNSKELCGNEVPPVPEQKQDTIISKILV 903
Qy 852 KAGSIEKYSQSILDLTDNRLQSTL-----KTLSPFPPTVTVGSEVOITTAIDVLP 906
Db 904 EPEGLEH-----EVTFNLSLCPMGAEVSELALALKPDSVVEBSARASTVGLDILGS 955
Qy 907 SINGLASLIRMPYCGEONMINPAPNIYILDTLTKKQOLTNLKEKALSEFKQGYOELL 966
Db 956 AMQNTODLLKMPYCGEONMVLFPAPNIYVLDYLNETOOLTOEIKTKAIAYLNTGYOROLN 1015
Qy 967 YOREDGSPSARFNVDPSG---STWLSAFVLRGFLADPYDIDONVLRHTYTWLKGHOX 1022
Db 1016 YKRDGSGISARF--DKPRNHANTWLTAFVLSKSPAQARKYTFIDEVHITQALLMSQOOK 1073
Qy 1023 SNGEFMPDGRVYHSELQCGNKSPTLTAYIVTSLGKRYQPNIDVQESILFSEF--- 1079
Db 1074 DNGCFRRSGSLNNAMKGVDEVTLSAYITILLSESLPTHVNVNNALEFCLDTANKSA 1133
Qy 1080 ---SNGISDNTYLLALITLASSVSGPKA--KEALNMLTWRBEOGGMQFVWSESLSDS-- 1134
Db 1134 RGAGGSHVYTKALLAYAFALAGNODTKEIKSLIDEAIVEEDSVMTWRPQ--RPSVSVG 1192
Qy 1135 --WQPR--SLDIEVAAYVALLSHFLOFQTSSEIP-----IMRWLSRQRNLSGFPAS 1180
Db 1193 LMYOPQATSAEVEHTAYVLLAYL---TTEPAPQOEDITTAAMLIVKMLTYKQONSHGGRSS 1248
Qy 1181 TODTVALKALSER--AALMTERNTNIQVTVGPSSPSBLAVVQ-----PM 1224
Db 1249 TODTVALHAKSKGSAFTPAKKAQVITRSSGTFKPGVNNNNOLLQRTVLPVPG 1308
Qy 1225 AVNISANGFPAICQOLNVVNVKASGSSRRRSRIONQDAFPLDVAVK---ENKODLNHV 1280
Db 1309 DYTAKVTVGEGCVVLTQTSIKYSV-----LPREEFPFAVVVOTLPTGTCEDPRAHT 1357

```

```

Qy 1281 DLNACTSPSGPG---RSGMALMEVNLISGFWPSEASISSET---VKQVEDHDKLNTYL 1334
Db 1358 SFOISLINSITYGSRSESNMAIADVKNVSGRIFPLKPTKMLERSVHSRTEVSNHHVILYL 1417
Qy 1335 DSVNETOFCVNIPAVNRFKVSNTOASVSIVDYEPERRQAVRSYNS 1380
Db 1418 DKVSNQTVNLISFTVQODPIRDLKPAVKVYDYERKDEFAVAKYSA 1463

RESULT 4
A2M1_MOUSE
ID A2M1_MOUSE STANDARD; PRT; 1476 AA.
AC P28665;
DT 01-DEC-1992 (Rel. 24, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Muringlobulin 1 precursor (Mug1).
GN MUG1 OR MUG-1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 28-57.
RC TISSUE=Liver;
RX MEDLINE=91358495; PubMed=1840592;
RA Overbergh L., Torrekens S., van Leuven F., van den Bergh H.;
RT "Molecular characterization of the muringlobulins.";
RL J. Biol. Chem. 266:16903-16910 (1991).
CC -1- FUNCTION: A PROTEINASE ACTIVATES THE INHIBITOR BY SPECIFIC
CC PROLYSOLYSIS IN THE BAIT REGION, WHICH, BY AN UNKNOWN MECHANISM
CC LEADS TO REACTION AT THE CYSTEINYL-GLUTAMYL INTERNAL THIOLESTER
CC SITE AND TO A CONFORMATIONAL CHANGE, WHEREBY THE PROTEINASE IS
CC TRAPPED AND/OR COVALENTLY BOUND TO THE INHIBITOR. WHILE IN THE
CC TETRAMERIC PROTEINASE INHIBITORS STERIC INHIBITION IS SUFFICIENTLY
CC STRONG, MONOMERIC FORMS NEED A COVALENT LINKAGE BETWEEN THE
CC ACTIVATED GLUTAMYL RESIDUE OF THE ORIGINAL THIOLESTER AND A
CC TERMINAL AMINO GROUP OF A LYSINE OR ANOTHER NUCLEOPHILIC GROUP ON
CC THE PROTEINASE, FOR INHIBITION TO BE EFFECTIVE.
CC -1- SUBUNIT: Monomer.
CC -1- TISSUE SPECIFICITY: Plasma.
CC -1- SIMILARITY: TO OTHER PROTEINS OF THE ALPHA-MACROGLOBULIN FAMILY,
CC INCLUDING COMPLEMENT COMPONENTS C3, C4, AND C5.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.1sb-sib.ch/announce/
CC or send an email to license@1sb-sib.ch).
CC -----
DR EMBL; M65736; AAA73048.1; -.
DR HSP; P01023; 1BV8.
DR MGD; MGI:99837; Mug1.
DR InterPro: IPR002890; A2M N.
DR InterPro: IPR001599; MacrogloblnA2.
DR Pfam; PF00207; A2M; 1.
DR Pfam; PF01835; A2M_N; 1.
DR PROSITE; PS00477; ALPHA_2_MACROGLOBULIN; 1.
KW Serine protease inhibitor; Glycoprotein; Plasma; Bait region; Signal;
KW Multigene family; Thioester bond.
FT SIGNAL 1 27
FT CHAIN 28 1476 MURINGLOBULIN 1.
FT DOMAIN 677 734 BAIT REGION (APPROXIMATELY).
FT DISULFID 48 86 BY SIMILARITY.
FT DISULFID 251 276 BY SIMILARITY.
FT DISULFID 269 288 BY SIMILARITY.
FT DISULFID 461 555 BY SIMILARITY.
FT DISULFID 587 773 BY SIMILARITY.
FT DISULFID 634 680 BY SIMILARITY.
FT DISULFID 849 885 BY SIMILARITY.

```



QY	686	-----	VRGHPEETIMTLDINMGYRIYOEFTVYVDPDSITSW	720
Db	708	RTTSLGRPEEPKROBSSNDPTEIRIKRPFETVMQDLYVNSGLAEVEMTPDITTEM	767	
QY	721	VATFVISEDGLGLTTTPVELQAFOPPFIFLNPYSYIRGEFEALEITTFNLYKDATEV	780	
Db	768	KGALCLSDTGTGLGSSV-VPLQAFKPEFEVSLPYSVIRGEAFMLKATVMNLYPTSMQ	826	
QY	781	KVIEKSDKFDLMTSSSEINATGHOQTLTVSEDEGATVLPFIRPTHGEIPITYTALSPT	840	
Db	827	SVGLEASPDFTAVPVGDD-----QDSYCLSANGRHRTSMVLTYSKSGANNFVSAAQ	880	
QY	841	AS-----DAVTOMILVAEISIEKYSQSILDLTDNRLOSTLKTISFSF	884	
Db	881	SSPFGSEVATVPATGRKDTYVKULIYVEBEIKGEHFFSSIFCASDAEIS---EKSSSGP	937	
QY	885	PPNTVTSRROVITAIIGDVLGPSINGLASLIRMPYGCGBQMINFAPNIIYLDLYTKKQ	944	
Db	938	PPFVVKDSARAFHSVGDILSSAIRNTQNPPLHMPYGCGBQNMVFAFNIIYVLKYLMETQ	997	
QY	945	LTDNLEKRLSMRQGYRELLYQREDSPEAFG---NYDPSGSTWLSAPVLRGFLDEPV	100	
Db	998	L7QKTKTKALGLRAGYORELNYKHKDSYSAFEDQGBERGNTWMLTAPLAKSPAARAF	105	
QY	1003	IDIDQNVLRRTYTMKGHQKSGEFPWDEGRVYIHESELQGNKNSPVTLLAYIVTSLIGRYK	106	
Db	1058	IFIDESHIHTATWLSQKQKONGCRRSGSLFNNAKMGQVDDEHTLSAYITMALBESSLP	111	
QY	1063	QPNIDVQESIHFLS-----EFSRGISDNYTLALITYALS SVSGP -KAKBALNMLTWRAE	111	
Db	1118	ATHPVVSKALSCLESSMKTIBOERBASFYVTKALMAVAFALAGNQNRDEILKSDBEAL	117	
QY	1117	QEBGMQFWV-SSESKLSDSW-----QPRLDIEVAAYALLSHLELOFQSEBIP-----	116	
Db	1178	KENNSITHMRPKSKSRKSHHLYKPPASSAEVEMNAVYVLAEL---TAQPAPEPEDITLS	123	
QY	1164	--IMRWLSQRNLSLGGFASTODPTVYALKALSEFPAAL---MNTERTNIQVWYGPSSP---	121	
Db	1234	MSITMMLTYQGNNSNGFSSSTQDTVALDALSKYGAIVFBSRQKTLTYTIGSTGSFQKQ	129	
QY	1216	---SPLAVVOQNAV-----NISANGFPAICQNLNVNYKASGSSRRRRRSIONOEAFD	126	
Db	1294	VENSNNLLIQVALPDIPEDYTIISVSGGCYVQAQTMRLRYNMHLE-----KQLSAFA	134	
QY	1266	LDAVAK-----ENKDDLNVDLNVCSSPEFG-GRGGMALMEVNLISGFMVSEALSSETV	132	
Db	1345	IWQTVPLTCNNPKGNSFQISLEISYSGSRPANNMVIADVKMLSGF-----IPLKPTV	139	
QY	1321	KKV-----EYDHGKLNLYLDSVNETQCVAIPAVRPFKVANTODASVSIYDYEPR	137	
Db	1399	KLERLERLEHRSREVENNNVLIYLOQVNTQTLAFSEFIQQDIPRNLQPALVXYDYIETD	145	
QY	1372	ROAVRSYNS	1380	
Db	1459	EMAFAYSS	1467	
RESULT 5				
A2M2_MOUSE				
AC	P28666;	STANDARD;	PRT;	1451 AA.
DT	01-DEC-1992 (Rel. 24, Created)			
DT	01-DEC-1992 (Rel. 24, Last sequence update)			
DT	15-SEP-2003 (Rel. 42, Last annotation update)			
DE	Murtaglobulin 2 precursor (Mug2).			
GN	MUG2 OR MUG-2.			
OS	Mus musculus (Mouse).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			
OX	NCBI_TaxID=10090;			
NP	(1)			
SEQUENCE FROM N.A.				





FT	CONFLICT	691	691	N -> M (IN REF. 3)
Pos	CONFLICT	753	753	N -> O (IN REF. 2)
SO	SEQUENCE	1482	AA; 163836	MM; 0173EE61D2FC3B CRC64;
Query Match		19.0%;	Score 1396.5; DB 1;	Length 1482;
Best Local Similarity		28.3%;	Pred. No. 1.4e-68;	
Matches	432;	Conservative	280;	Mismatches 587; Indels 225; Gaps 53
QY	12	LLCVCTAALAAVAGPREFLVTA	PGIIRPGGNVTIGVELLHCHSPQVTYKAE	LKTSANLTV 71
DB	14	LLILLASADSNSTBEQYMWLV	PSLLHTEAPKK-GCVLLSHLNETVTSASLESGRNRS-	71
QY	72	SVLBAEGVPEKSFK--TLT	PSLINSADLEYELRVGTRTODELIFS	STRLSFEYTKR 129
DB	72	--LFTDLVAEKOLFVCSFT	LPRIASASSVAFLSISIKRPTD--FRKNTV	LVLTQSD 127
QY	130	SVFIQTDALAYKPKQEVKFR	IVTLFSDFPKYSLNIL-IKDPKSNL	IQOMLSQOGLV 188
DB	128	-VFQGTDKPMYKRGQTVR	FRVVSVDENFPRHMLFLIYLBNRRNRIR	IQOMSLKLEAGI 186
QY	189	ISKTFQLSHPILGWSIQOV	NDQTYV-SFQVSEYVLPKFEVTLQ	TPLYGSMNSKHIL 247
DB	187	NQLSPFLSEPIGSGYRVV	QTESGRIQHPFVEEFVLPKEVKQV	PKIISIMDEKYN 246
QY	248	GTITAKYTYGKRVKQDVT	LTPLPLSPWKKKNI	TKTFKINGSANSEFNDEMKNNVDSN 307
DB	247	ITVCGEYTYGKRVPGLAT--	VSICRLSRNLNCDKQVCEEFSS--QQL--	--NSN 293
QY	308	GL-----SEYDLSPPG--	PVEILTYTESVTGI--SRNVS--TNV	PFKOHDIIEF 353
DB	294	GCITQGVHTKMLQITNTG	PEMKLRVAKRIREBTDLEVTANR	ISLETNIVSK-----LKF 348
QY	354	PDYTTVLKPSLNFATV	KVTRADGNQLTLEERNNVITV	TQNTYEYWSGNSG----- 408
DB	349	VKVDSHFQGIPEFFA--	QVLVADGKGVPIPNKL--FFISV	NDANYNSNATTNEQGLAQFS 404
QY	409	-NQRKEAVQKI--NTVP	QSGTFKIEFILLDSSELQ	KATFLGSKSSMAVHSLFKSFS 464
DB	405	INTTISIVNKLFRVFTV	HPNLCPHYSW--VAEDHQAQ-----	HTNRVPSL----S 451
QY	465	KTYIQLKTRDENIKVQ-----	SPFELVYVSGNKRILKELS--YVVS	RSGOLVAVG----- 510
DB	452	GSYIHLEPVAGTLPGCH	ETTITTAHTLNRQMGELSELSPHIL	IMAKGIVASGHTTLEY 511
QY	511	----KQNSTMFSLTPENS	WTPRACVIVYIIEBDEIISDVL	KIPVQVFNKIKLWYSKV 566
DB	512	ESGDMKGSFALSPVES	DVAPFIARMPITFALPDGEV	VGSGSEKREIENCLANKVDLSFSPA 571
QY	567	KAEPEBKSLRISVTPQ	PDPSIVIGIVAVDKSVNIMANSNDIT	MENVVHLE----- 615
DB	572	QSPASHAHLQVAAA--	POSTCALRYAVDOSVLLMKPEAL	SVSVNLLTVVKDLTNEPDNV 630
QY	616	-----LYNTGYVILGEM	PMNNS---FAVROEGCLMWLT	PDANTLKQYIDGYDN 657
DB	631	DQOEEOGHCPPEFTIH	NCALIVPLSSNEADYISFLKMG	LKVFTNSKRIRPKSGCVIRS 690
QY	658	AE-----YAEFMEENEG--	HIYDI--HDFLSGSPH--VRKH	FPETWIMLDT 699
DB	691	VSAGAVGGYVAGLG	VERPRVYVPOQTGNVPI	PLNNEGSGPRPELVRSGFPETWIMELY 750
QY	700	NMGKIVLYOFEVTV	PDSTISWATGIVSEDLGLTTTPVEL	QAOQFFIFLNLPSYVI 759
DB	751	AVNSSGVAVEGTV	PDITITEMKAGAFCLSEBDAGL	GISST--ASIRAPQPFVELTWPYSVI 809
QY	760	RGEERALETITN	YKDADEVVILIEKSKPDLMTSS	HNMTGHQDTLLVSEBDATVL 819
DB	810	RGEVTLKATVANTY	PKCIRVSVQLKASPAF-----	LASQNTKGESYICGSE--RQTLS 863
QY	820	PIPIRETHIGE	IRITVYALSPTA-----SDAV	OTMLVVAAGIEKSYSOS 863
DB	864	WTVPKTLTG	NVVSVAEMQSLBELCQNEV	VEPEIKRQDTYIKILLVBAEGIEQKITS 923
QY	864	ILLDLITDNR	LOSTLKTLSFSPPNVTY	GSERVOITAIQDVLGDS INGLASLIRMPYCGCE 923

```

Db      924  ---SMTCASGANNVSEQLSLKPSNVVESHARASFSVLGDIIGSAMQNIQVLLGMIPYCGE 980
Qy      924  QNMINPAPNIYILDYLTKKQLIDNLKEALSEFKOQYRELLYQREDGFFSAFG-NYPD 982
Db      981  QNMVLPAPNIYVLYNLNETQOLTOEIKAKKAVGLITYQROLNKHODGSAFSTGERGYR 1040
Qy      983  S-SGTTWLSAPFLRFLPADPYIDIDQNVLRRTYTWLKGHKSNEFMPDPVRVHSELQGG 1041
Db      1041  NQGMVTLTAFLPFLKFAARSAYSIFIDEAHITQSLTWLTSQMKDNCGFRSSGLNNAIKGG 1100
Qy      1042  NKSPTVTLTAAYIVTSLGKRYKQPNIDQESIHFLFS-----EFSRGISDNYTLALITYA 1095
Db      1101  VEDATISAVYTTALTBLEIPLEPVNPIRYNMLPCLSESAMNVAKECTHG-SHYTTALLAYA 1159
Qy      1096  LSSVGSPP-KAKEALNMLTWRAEDGEGMQFVWSSEKLSD-----SWQPSRIDFVAAYAL 1149
Db      1160  PSILGKQNRREILNLSIDKEAVKEDNLTVMHERPQRPRAPGHLXOTQAPSALEVMTSYVL 1219
Qy      1150  LSHFLQVQTSNGIF-----IMWLSRQRSLGSGFASDTDTYALKALSEF--AL 1197
Db      1220  LAYI---TAQAPATSGDLTSATNIVKIMKQNAQGSFSTQDTVALHALSLRYGAATF 1275
Qy      1198  MNTERTNIOYTVVTPSPSSPSP-----LAVQPMAYN-----ISANGFGAICOLNV 1242
Db      1276  TRTKKT-AQVTVQDSQFTSTNFOVDNNNLLILQOISLPELGEVITYVTGERCYLQTSM 1334
Qy      1243  VYNYKASGSSRRRRSISNOEAFDLDVAVKENKODLN--HYDLNVCTSSFGPGR--SGMA 1297
Db      1335  KYNLIPE-----KEDSPFALKVQTVTPQTCGHKRAHTSFQSLTISTYGNRPASNMV 1385
Qy      1298  LMEVNLISGFWNVSBAISL---SETTKVGEYDHCKMLYIDSVNETQPCVNIAPVRKFX 1354
Db      1386  IVDYKVMVSGFPLPKPTVKMLERSSVSRTESNNHVLITYEQVNTQTLSSFMVLQDIPV 1445
Qy      1355  SNTODASVSIYDYEPYRRQAVRSY 1378
Db      1446  GDLKPAIVKVIYDIETDESVAET 1469

RESULT 7
OVOS_CHICK
ID_OVOS_CHICK  STANDARD;  PRT;  1473  AA.
AC  P20740;
DT  01-FEB-1991 (Rel. 17, Created)
DT  01-NOV-1995 (Rel. 32, Last sequence update)
DE  01-NOV-1995 (Rel. 32, Last annotation update)
DE  Ovostatin precursor (Ovomacroglobulin).
OS  Gallus gallus (chicken).
OC  Eukaryota; Metazoa; Chordata; Craniota; Vertebrata; Euteleostomi;
OC  Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC  Gallus.
CX  NCBI_TaxID=9031;
EN  [1]
RP  SEQUENCE FROM N.A.
RC  TISSUE=Oviduct;
RA  MEDLINE=95218210; PubMed=7535598;
RA  Nielsen K.L., Sotttrup-Jensen L., Nagase H., Thøgersen H.C.,
RA  Eizerodt M.;
RT  "Amino acid sequence of hen ovomacroglobulin (ovostatin) deduced from
RT  cDNA.";
RL  DNA Seq. 5:111-119(1994).
RN  [2]
RP  SEQUENCE OF 37-49.
RC  TISSUE=Egg white;
RA  MEDLINE=83238315; PubMed=6408074;
RA  Nagase H., Harris E.D. Jr., Woessner J.F., Brew K.;
RT  "Ovostatin: a novel proteinase inhibitor from chicken egg white. I.
RT  Purification, physicochemical properties, and tissue distribution of
RT  ovostatin.";
RL  J. Biol. Chem. 258:7481-7489(1983).
RN  [3]
RP  SEQUENCE OF 976-1028.

```

CC TISSUE=Egg white; PubMed=7660577;  
 CC MEDLINE=93192299; Sottrup-Jensen L.;  
 CC Nielsen K.L., Sottrup-Jensen L.;  
 CC "Evidence from sequence analysis that hen egg-white ovomacroglobulin  
 CC (ovostatin) is devoid of an internal beta-Cys-gamma-Glu thiol  
 CC ester.";  
 CC Biochim. Acta 1162:230-232(1993).  
 CC -1- FUNCTION: IS ABLE TO INHIBIT ALL FOUR CLASSES OF PROTEINASES BY A  
 CC UNIQUE "TRAPPING" MECHANISM. THIS PROTEIN HAS A PEPTIDE STRETCH,  
 CC CALLED THE 'BAIT REGION' WHICH CONTAINS SPECIFIC CLEAVAGE SITES  
 CC FOR DIFFERENT PROTEINASES. WHEN A PROTEINASE CLEAVES THE BAIT  
 CC REGION, A CONFORMATIONAL CHANGE IS INDUCED IN THE PROTEIN WHICH  
 CC TRAPS THE PROTEINASE. THE ENTRAPPED ENZYME REMAINS ACTIVE AGAINST  
 CC LOW MOLECULAR WEIGHT SUBSTRATES (ACTIVITY AGAINST HIGH MOLECULAR  
 CC WEIGHT SUBSTRATES IS GREATLY REDUCED).  
 CC -1- SUBUNIT: HOMOTETRAMER, WHICH CONSISTS OF TWO PAIRS OF DISULFIDE-  
 CC LINKED CHAINS.  
 CC -1- PTM: THIS PROTEIN LACKS THE THIOESTER BOND FOUND IN OTHER MEMBERS  
 CC OF THIS FAMILY.  
 CC -1- PTM: CONTAINS 56 MOL GLUCOSAMINE PER MOL SUBUNIT.  
 CC -1- SIMILARITY: TO OTHER PROTEINS OF THE ALPHA-MACROGLOBULIN FAMILY,  
 CC INCLUDING COMPLEMENT COMPONENTS C3, C4, AND C5.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC -----  
 CC EMBL; X78801; CAA55384.1; -  
 CC EMBL; X78801; CAA55385.1; ALT\_INIT.  
 CC PIR; I50671; A20872.  
 CC HSSP; P01023; 1B98.  
 CC GO; GO:001711; F:wide-spectrum protease inhibitor activity; NAS.  
 CC InterPro; IPR002890; A2M\_N.  
 CC DR InterPro; IPR001599; MacroglublnA2.  
 CC DR Pfam; PF00207; A2M; 1.  
 CC Pfam; PF01835; A2M\_N; 1.  
 CC PROSITE; PS00477; ALPHA\_2\_MACROGLOBULIN, FALSE NEG.  
 CC DR Serine protease inhibitor; Glycoprotein; Bait region; signal.  
 CC KW SIGNAL 1 36  
 CC FT CHAIN 37 1473  
 CC FT CARBOHYD 67 1473  
 CC FT CARBOHYD 82 82  
 CC FT CARBOHYD 89 89  
 CC FT CARBOHYD 191 191  
 CC FT CARBOHYD 342 342  
 CC FT CARBOHYD 403 403  
 CC FT CARBOHYD 527 527  
 CC FT CARBOHYD 568 568  
 CC FT CARBOHYD 757 757  
 CC FT CARBOHYD 1141 1141  
 CC FT CARBOHYD 1221 1221  
 CC FT CARBOHYD 1315 1315  
 CC FT CARBOHYD 1347 1347  
 CC FT CONFLICT 46 46  
 CC SQ SEQUENCE 1473 AA; 166354 MW; A33C6847A14179BF CRC64;  
 CC  
 CC Query Match 19.0%; Score 1395; DB 1; Length 1473;  
 CC Best Local Similarity 27.0%; Pred. No. 1,6e-68;  
 CC Matches 421; Conservative 309; Mismatches 522; Indels 308; Gaps 59;  
 CC  
 CC 18 AALAVAPGRPLVTAIGIIRPGAVTIGVLEHSPQYVAELIKTASNLTVS---VL 74  
 CC 33 AAAGKEPQVYLVMPAVLQ-----SDSPSQVCLQFPMV---NQTISVRVYL 76  
 CC 75 EAE-----GVFEKSGKRTILTPSL-----PLNSADELY-ELRVTGRTQD-----ELIFSNS 119  
 CC 77 EYDTNTITTEFKNTTSSNGLQCLNMPVPTVSLSAFISFTAKGTTFDKERRSVIWMW 136  
 CC 120 TFLSEFTRKISVFIOTDKALYKPKQEVKPRIVTLFSDFKPKYKTSINIL- IKDPKSNLIQO 178

DB 137 E--SF-----VFVQTDKPIYKPGOSVWFRVVALDFNFKPQVEMYPPLAVQDPQNNRIQO 188  
 QY 179 WISQOSDLGVISKTQPSLSHPILGPMSTQV-QVNDQTYGQSVQSEYVLKREYTLQPL 237  
 DB 189 MONVTSEINIVIEEPLTEEPILGNVYKIIIVTKSGERTSHSLVEEYVLPKEDVITVTAG 248  
 QY 238 YSNMSKHLNGITTKATYTGKPKVDVTLTFL-PLSPFGKKK-----NITPKFINGS 290  
 DB 249 SLTWDSSELTVYICAVYITYGQVEKQVLSVARDSDGRCKSKSVCSFTDLDTDGL 308  
 QY 291 NFSFND--EMKNVWDSNGISEYLDLSSPGVEILITVTESVTSISRVNSTVFPKQHD 348  
 DB 309 SHILSKVFEINRI-----GYRKNLD-----VKRAIVTEKEQVNLTAQGISI--TQVM 355  
 QY 349 YIEFPDVTYVLKPLNFTATYKVTYRADGNQTLTEERNNVITYTQRYTEYWGSGNSG 408  
 DB 356 SLOFENVDHNYRGIPIFGQIKLVKDKNSP-----SNKYQLFVNKNKTH----- 402  
 QY 409 NQKMAVQKINTVPOSG--FKIEPILDSSELQKAYFLGS-----KS 452  
 DB 403 -----NFTDINGIAPSIDTSKIFD-PELSLALYKTSQCHSEGTIEPSYDPA 451  
 QY 453 SNAVSLKPSPKTYIOLKTRDENIKVSP-----FEIIVSGNRLKELS--YMWVSRG 504  
 DB 452 SLTVQRLV-SWTSFVRLEPLMKDSCGQKMITVYIILNTEGHEHINIVNYYGMKAG 510  
 QY 505 QLVAVK-----QNST-WFSLTPENSWPKACVITYIYEDDEIISDVLIKIPVOLVF 555  
 DB 511 KIVLGEIKVNIQADQNGTFMPLVNVNKKMAPALILVYVMPAKELVADSVRFSEIKCF 570  
 QY 556 KKKIKLWYSKYAEBSEKVSRLISTVQPSIVGIYAVDQSVNLMAANDITMENV--VHE 613  
 DB 571 KKKVOLQSEKMLTTSNVSLVLEBA-ANSPCAVAVDQSKMLLSSETLSSETTYNHP 629  
 QY 614 L-----ELVNTGYVL-----GWFNNSFAVFOEGMLVLD 643  
 DB 630 IODLOGYIFNGLNLEDDQDPCVSSDDIFHKGILYRPLTSLGPRVYQELRMGKKFTFN 689  
 QY 644 ANL-----TKDYI-----DGYYNDAEYAEERMEENEGHIVDHFSLGS 682  
 DB 630 SKIROPTVCTRETPVPSFYFLNAGFTASTHVKLSAEVARE--ERKSHILLET----- 740  
 QY 683 SPHVAKHPETMIW---LDTNMGYIYQEFVTVGDSITTSWATGTVFVISEDGLGLTTP 739  
 DB 741 --HEFPEETWIMDIILINSTG---KASVSTYIPDTITTEWASAFVCEELAGFGM-SVP 793  
 QY 740 VELAQFQFFFLFPLNPFYVIRGEEFALETTFNYLKQDAVEVUIIEKSDKFDILTSSBI 799  
 DB 794 ATLTPQFPPVDLTPYISIIHGEDFLVAVANFNVYLNHCIKINVL-----L 839  
 QY 800 NATGQOQLVLVSEDEGA-----TVLPPIRPHLGEIPIVTVA-----L 837  
 DB 840 ESLDVAQLISPEDGCACAKIRKSYVNNIPKRGDVLFSITATNDDEACEEALANI 899  
 QY 838 SPTADAVTQMLVABEIEKSYQSILDLTDNRLOSTLKLSSFPNPTTGERVOI 897  
 DB 900 RIDYVDYQIRLVLVEBIEIRRETONFLICMKDVIS--DPAIDLTNNVVEGSPRSPF 956  
 QY 898 TAIGDVLPSINGLASLIRMPYGCGEOMINPAPVITYLDTYTKKQPLTDNKEALAS 957  
 DB 957 SVYGDIMGTALIONVQLOMPFGNGBONNVLPAPVITYLDTYTKKQPLTDNKEALAS 1016  
 QY 958 ROGYORELLYQREDSFSAFGNYDPSGTSWLSAFVLRGFLVLEADPYTIDONVLRHTYWL 1017  
 DB 1017 VSGYQKQSYKHPDQSTFGIRBDEGNTWLTAFYFKSFEASRFTYIDDNVQAQTLWL 1076  
 QY 1018 KGHQKSNFMDPRGVHISELOGKNSPVYTLATVTSLLGTRKQNPINDQESHPLES 1077  
 DB 1077 ATKQKTDGCFQSTGILVNNANKGAVENELSLSAVITTAALLGHSMTVIRNAPCYLET 1136  
 QY 1078 EFSRGISNNTYLLATTVYALSSVSGPKAKEA-LNMLTWRAEGBGQGFVWSSE-----SK 1130





```

QY 243 SKHNTITAKTYGKPVYGDVTLTFLPLSPFGKKKNTKTFKINGSANFEND---EE 298
Db 241 EELPTTACGVYTKVPGLVTL-----RVCKKYSRSTGCHNOMSMISICAB 288
QY 299 MKGVNDSNGLSLEYL-----DLSSPG---PVELTIVTE-----SYTGISRNVSTVFPK 345
Db 289 FSGQADKCCFSGVYATKVFQLSQKHDKIEVAKIKKEGIGIEITGIGSCETIANALSK 348
QY 346 QHDYIIIEFDYTTVTLKPSINFTATVYKTRADGNQLTEERRNNVTVTQRYNTEYSGS 405
Db 349 -----LKFVKVNTNRPGLPFGSQVLLVDEKGPFP-----NKQITSVSPGLGLSLFTTD 399
QY 406 NSGNQKMEAVOKINTVYPOSSTFKIEPFLLEDSSSELQKAYFSGSSMAVHS--LEKPD 463
Db 400 EHGILNT-SIDTSNFTAP-----FLRVVTVYKONHCYDNMMWIDEFHTQADHSATLVFSP 453
QY 464 SKTYIOLKT-----RDENIKVSGPFELVYSGNRKLELSRMVVSROQLAVGKONST 515
Db 454 SOSYIQLLELVFGTLACGQTOEIRIHLYLMEDIMKNKTLTFEYLKARG---SIGMLGSH 510
QY 516 MFSLEPEN-----SWTPKACVIVYIEPDGELISDVLKIPQVLFKXIKLY 562
Db 511 VLSLEGKMKGVFSLPIQVEPGMAPRAQLITAILNEBELVADQNFEIEKCFANKVNS 570
QY 563 WSKVYAEPSKVSLSRISVTQPDISVIGIVAVDKSVNL-----599
Db 571 FPSAQSLPASDTHLKVKAA-PLSICALTVADQSVLLKPEAKLPSQISVNLPGKTVQGA 629
QY 600 -----NASNDITMENNVHELINTGYIIGMFANSRAVPECGMLVLTDAHLTK--- 648
Db 630 FPGVPYKCHENCITSGEDITHNGIIVTPKHSIG-DMDAHSIFQSVGINIFTMSKIKRPF 688
QY 649 -----DYIDGYVDA-----EYARFMEENEGHVDIHDS 679
Db 689 CQEPQHYPMGVAPALAAVAASSGSSSRAMGVPMAGDYSER---NOVREPR--- 741
QY 680 IGSSPHVRGHPETWIM---LDTNMGYRIYQEFVTVPSITSWATGFSVISEDGLGI 735
Db 742 -----TVRKYPETWIMDLVPLDVSGD-----GELAVKVDPTITEMKASAFCSGTTGLS 792
QY 736 TTTTPELOAFOPFFITLNPYSYIRGEERALETTINYLKDATEVYIIIEKSDKPIILMT 795
Db 793 SST-ISLQAFOPFFELTLTPYSVVRGEAFTLKATVNMVSHCOIIVDLISPDF----- 846
QY 796 SSEINATGHOOTLVSESDGATLPIRPHLGEIPIYATASPTA-----841
Db 847 -LAIVPGGHENSHCIGNERKTVSMVATPKSLGEVFTFAELAESQELCGNKLTVPAL 905
QY 842 --SDAVTQMLVYKAEIGIEKYSQSILLDLTNRLQSTLTKLSFSPPTVYTGSEVOITA 899
Db 906 VAKDITVVKSVIIEPEGIEKEQYNTLTCQDTELQD---NSSLELPPNVVSGSARATHAV 962
QY 900 TQDVIGPSINGLASLIRMPYGCGBQMINFAPNITYILDYITKKQULTDMLKEKALSFMRQ 959
Db 963 LQDITIGSAAQNLQNLQMPYGCGBQNMVLFVPNIYVLYNMYNETQULTEAIKSAINYLS 1022
QY 960 GYORELTYOREDSFSAFGNY---DSSGSTWLSAFLRCFLEADPYIDIDQNLHTHTYTM 1016
Db 1023 GYORQNLNYHSDGSYSTFGNHGGGNTPGNTWLLTAFLAKPAQOQSHIFLEKHTITNAFWM 1082
QY 1017 LKHQKNGEFMDPGVHISELOGKNSPVTLTAVYVTLGLGRKQPNIDVQESIHFE 1076
Db 1083 LEMKQKENCFOOSGVLNANNAMGVDDDEVTLTSAVITIALLEHPLVTHSAVRNALFCHE 1142
QY 1077 SEPSRGISN-----YTLALITYALSSVGS-PRAKALNMLTWRADEGGMQFVGS--- 1126
Db 1143 TAMA-SISOSOSHVYTKALLAVAFALACNKAARSELLESINKDAVKEEESLHWQPGDV 1201
QY 1127 SESKLSDSWQPR--SIDIEVAALSLHLPQPT-----SEGPIKMWLRSRON 1173
Db 1202 QKKKALSPQPRAPSVENTAVALLAVYLTSESSRPTRLDISSDLSSTASKIVYMIKQON 1261
QY 1174 SLGGFASQDPTVALKALSEFALAMTERTNIGVTVTGPS-----SPSPPLAVVQ 1222

```

```

Db 1262 SOGGLLTQDVTVALQALSKYGSATFT-RSQKEVLTSSSGTFSKTFHVNSGNRLLOE 1320
QY 1223 PMAVNISAN-----GFGPAILQNLVNVNKKASGSSRRRSRISQ-NOEARDLDAVAKND 1275
Db 1321 VALPDLPGVYVTKGSSGCVYIQTSLKYNILPYADKAPAPALQVNTLPLNFDA-----E 1375
QY 1276 DLNHDVLANCTSPSGP-GRSGNALMEVNLISGF--MVPS-EA1SLSEYKVEYDHQKLN 1331
Db 1376 DHRTFQIRLNVSYTGERPSSNNVYIDVKWVSGFIPKPSVKGLODQDPNIQRTVEVNTNHYL 1435
QY 1332 LVLDVNETQFCVNIIPAVNPFVYSNTQDASVSIVDYIEPRQAVNSYS 1380
Db 1436 IYIEKLNTQTLGFSFAVEDIPIVKLKPAPIKVYDYETDEFTVEEYSA 1484

RESULT 9
CO3_LAMJA
ID CO3_LAMJA STANDARD, PRT, 1673 AA.
AC 000685;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Complement C3 precursor [Contains: C3A anaphylatoxin] (Fragment).
OS Lampetra japonica (Japanese lamprey) (Entosphenus japonicus).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Hyperoartia;
OX Petromyzontiformes; Petromyzontidae; Lethenteron.
NCBI_TaxID=94989;
RN (1)
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RX MEDLINE=92251197; PubMed=1578150;
RA Nonaka M., Takahashi M.;
RT "Complete complementary DNA sequence of the third component of
RT complement of lamprey. Implication for the evolution of thioester
RT containing proteins."
RL J. Immunol. 148:3290-3295 (1992).
CC - FUNCTION: C3 PLAYS A CENTRAL ROLE IN THE ACTIVATION OF THE
CC COMPLEMENT SYSTEM. AFTER ACTIVATION (C3B), IT CAN BIND COVALENTLY,
CC VIA ITS REACTIVE THIOLESTER, TO CELL SURFACE CARBOHYDRATES OR
CC IMMUNE AGGREGATES. CYCLOSTOMATES C3 APPEARS TO REPRESENT THE
CC COMMON ANCESTOR OF MAMMALIAN C3 AND C4, SHOWING SIMILARITIES TO
CC BOTH PROTEINS.
CC - SIMILARITY: TO C4, C5 AND ALPHA-2-MACROGLOBULIN.
CC - SIMILARITY: Contains 1 anaphylatoxin-like domain.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
CC EMBL: D10687; BAA00983.1; -
CC PIR: I50806; I50806.
CC HSP: P01024; I03D.
CC InterPro: IPR002890; A2M_N.
CC InterPro: IPR000020; Anaphylatoxin.
CC InterPro: IPR001599; MacroglobinA2.
CC InterPro: IPR001134; Netrin_C.
CC Pfam: PF00207; A2M_N.1.
CC Pfam: PF01835; A2M_N'1.
CC Pfam: PF01821; ANATO.1.
CC Pfam: PF01759; NTR.1.
CC ProDom: PD003264; Anaphylatoxin.1.
CC SMART: SM00104; ANATO.1.
CC SMART: SM00643; C3A5C.1.
CC PROSITE: PS00477; ALPHA_2_MACROGLOBULIN.1.
CC PROSITE: PS01177; ANAPHYLATOXIN_1.1.
CC PROSITE: PS01178; ANAPHYLATOXIN_2.1.
CC Complement pathway; Plasma; Inflammatory response; Glycoprotein;

```

KW Signal; Thioester bond.  
 FT NON TER 1  
 FT SIGNAL 1  
 FT CHAIN 14 1673  
 FT CHAIN 14 653  
 FT CHAIN 1379 1673  
 FT CHAIN 1379 732  
 FT DOMAIN 678 713  
 FT DISULFID 678 705  
 FT DISULFID 692 712  
 FT DISULFID 692 713  
 FT CROSSLINK 986  
 FT  
 SQ SEQUENCE 1673 AA; 18767 MW; 485FC756B123A506 CRC64;  
 Query Match 14.0%; Score 1026.5; DB 1; Length 1673;  
 Best Local Similarity 23.8%; Pred. No. 3.1e-48;  
 Matches 385; Conservative 295; Mismatches 619; Indels 319; Gaps 63;  
 12 LACVCTAALVAPGPRFLVAPGRIIRPGWNTIGVELLECPSS---QVTVKAEILKTAS 67  
 1 VLLMSGISTVTQDPMVLISVPSVILIGSDVNV---LVDAHASTEDVSVVVRAEFEITKK 57  
 68 NLTVSVLAEGVPEKSGFKTLTLPSTL-----PLNSADEIYELRVYVGT 110  
 58 QLAT-----QTITLTQDLPALATLKLGFIDENPKTNASTKHHRLVAVY 103  
 111 QDELTPSNSTR--LSFETKRISVFIOTDKALYKQGVKFRIVTLFSDEK---PYKSL 164  
 104 ESKSFNKEITPAHALLSTRSGH--VVVQTDKPIYTPDEKVKYRMPNMRREVHRIPVQSM 161  
 165 NIKIKDKPSNLIOQWLS--QOSDPLGVSKTF---OLSHPIGLDMSIOVANDQVYOS 218  
 162 TVQIVADGVIVERQITIKATIDEGIVDGISFTIPAIKSK---GTWKIFAPMGAAPINS 218  
 219 ---FOVSEYVLPKREVTLOTPLYCSMNSK---HLNG---TITAKYTYGKPYKGDVTL 266  
 219 SAFEVDVEXILPTFEV-----KINPKQRFHINDEEFVVIDITANYFQNELVSGTAVY 270  
 267 T-FLPLSFNGKKKNIITKFKINGSANFSFDEKKNVMSDNGLSSELDLSSGQPVILT 325  
 271 RYFLENGADVPLVDSSSTTLVAGEGLSILKKEKLLKFPNAKDLAF-----SLTIKT 323  
 326 TVTESVTVGISRNV--TWVFFKODHYIEFFDYTVLPSLNFATVTVTRADGNQLLEE 384  
 324 TVLSSQAETBEAEVLGKIKIVESRYQTATKTSYKFPBELPYFIQVEVRNADGS---PS 379  
 385 RNNNVTVTVQRYVTEYSGNSNGNQKMEAVOKINTVTPQSGTFKLEFPILEDSELOLK 444  
 380 KEYDVAVKVGVSAT-----INPQKMTDSNGLTSPVTPPNVNLVTVTRDERH----- 430  
 445 AYFLGKSSNAVHSLFSPSKTYIQKTRDENTIKVSPFELVVSQN---KRLKELSYNV 500  
 431 ---PSNEQGEIYVYTAQKASASVYMHIDV--TRIMIGETLNVFLAKTQTOLNAVTHFVYV 486  
 501 VSRGOLVAV-----GKQNSTMFSLTPENSWTPKACVIVYIIEDEGELISDVLIKIPVQ 552  
 487 LITGVAVIKINRKTKESGGGSPSNVIRIPITPD--MAPRRFLAYVILPGELVLAOSVYVEVT 544  
 553 LVFNKNKIKLYMSKVAEPS--EKVSLRIS--VTPQDSIVGIVAADKSVNLNANSDITMEN 609  
 545 ELCKSQVSL---SLKGRPTLEPKAMLLDLIGERDAVGLLADVQAVYANRRKRLTQDR 601  
 610 VVHELELYNTGYLYGMFNSFAVFOECGLVLTDLANTLTKYIDGVYDNAVY----- 660  
 602 VMKAMEFTDGTCAEGAGRGVFSXGLALITSKGL-----NTDRSEIGCEPKVSRK 655  
 661 -----AERMEENEGHIVD-----IHPF----- 678  
 656 PROLSMLQIRREAEKTYQEFKCCVVDGLKMSPTGGQGEERLKRVTGKECVDAFLQCKK 715  
 679 -----SLGSSPHVAK-----HFPETWIW--LDTNNGVRIYQEF 710

DB 716 AEEYRKSESIGAKTVLRNDFMELDMNDEVMMAVFPQSMQMNKXKNSCKYGRHPQIR 775  
 QY 711 VTVPSDITSVAVANGFVISEDGLGTLTPPELOAPQPFIFLNPVSVIRGEEFALEITI 770  
 DB 776 LQDPDITTTNNQAVSISKRGVCL--ADPLLVTSTQDFLKLHLPYSVKRGEOTEIRVIL 834  
 QY 771 FNYLKDATEVKEVIEESDKDEDIAMTSSEINATGHQOTLVPSBEDGATVLEPIRPTHGE- 829  
 DB 835 YNMMSLSL---TLTMDIVESICSTSKGAKRSQGS--TYKGGANVVSPIPLKIGEH 890  
 QY 830 -IPITYTALSPTASDAVTOIILVKAQIEKYSQSLILDLTDNRLOSTKLTSFSPFPNT 888  
 DB 891 HISIRRVYGRTPGDGDVQKILRVAPEGVDIRSESRSVHAEERETFEIKNEIS----PDV 946  
 QY 889 VNSERVQITAI--GDVLGSPS-----NGLASIRMPYGGEGOMNPNANVIYLDVLT 941  
 DB 947 VPNSDVLTVFISVGDLEIETMNVNCLDAKSTSNLIQIPYCGEONMKMAFTTLLTYLDS 1006  
 QY 942 KKO---LTDNLKEKALSFRQGYQRELYQREDSFSAFQNYDPGSGTWLSAFVLRCFLE 998  
 DB 1007 VOEMEKIGLRREALIGFLKQYSRELSTYRKADHSAAF--IKRPSWTALFVAVKYSL 1064  
 QY 999 ADYIDIDQNVLRHTYTW--LKGHQKSNGEFMDPGRVHSELQ---GGNKSPTLVAVYV 1054  
 DB 1065 AKRVIIVDQELCGPWEIITKNNQSDGSYREDDGPVIRHMOQGVGTGEGVMTAFILI 1124  
 QY 1055 SLIGRYKY---QPNL--DVQESIHFLSEFSGISDNTLALITYAL-----SSVGS 1102  
 DB 1125 GIGQAGEYGVSVENKQSNRAVOFLASKVS--DLRMVITIALTRYALALQDESEAH 1183  
 QY 1103 KAKEALNMLTPRAEGCGMQFVWSSEKLSDSQPSLDIEVAVALLSHFLQFQTSSEGI 1162  
 DB 1184 SWKLENRITFEK---GHRVWKAET--SHVLRMAISIVEARVAGLTYLRKQDVESAR 1238  
 QY 1163 PIRKWLSPKQNSIGFASFTODITVALKALSEFAA--LANTERTNIQVTVGPS- 1214  
 DB 1239 EIVDMLEORNYGGGFQSTODITLALQAMAQYMDSSSKELLIDVQLEITSPPKNNFEK 1298  
 QY 1215 -----PSPLAVVPMANVISANSGFPAIQOLNVVNVKASGSSRRRSIONQAFDL 1266  
 DB 1299 ITEETRFVQEPHKIPGNGNITIKASGTFYTSIMSVFNKVAOSS-----KSCSTFDL 1351  
 QY 1267 DVAVKENKD-----DLNHYD-----LNVCTSFSG-----PGRSG 1295  
 DB 1352 KYMTAADGESQGLGFWDGKRRRDIGBGCVIAYVYRMKCTRYKPKEDLSSESG 1411  
 QY 1296 MALMEVNLISGFVWP--SEASISLSETVK---VEYDHGKLNLYLDSVNETQ--FCVNIPA 1348  
 DB 1412 MTIIEVNMLTGF--IPDKNDLILQKESVDKYISNYEITDSVLITLYMDKVPSTEDYCF 1470  
 QY 1349 VNFKYSNTQDASVSLVDYEPKRAVRSYN---SEVKLSS--C--DLCSVQGCRCCE 1400  
 DB 1471 KQMLRSDMIQPVYASVYDYSPADKCTRLYNLPGGYELSEPLCQNDLCCQVEVSCPAK 1528  
 RESULT 10  
 CO4\_HUMAN STANDARD; PRT; 1744 AA.  
 ID CO4\_HUMAN P01028; Q13160; Q13906; Q14835; Q9NPK5; Q9UIP5;  
 AC P01028; Q13160; Q13906; Q14835; Q9NPK5; Q9UIP5;  
 DT 21-JUL-1986 (Rel. 01, Created)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DT 15-SEP-2003 (Rel. 42, Last annotation update)  
 DE Complement C4 precursor [Contains: C4a anaphylatoxin].  
 GN C4A AND C4B.  
 OS Homo sapiens (human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A. (C4A AND C4B).  
 RC TISSUE=Liver;  
 MEDLINE=84156544; PubMed=6546707;

RA Belt K.T., Carroll M.C., Porter R.R.;  
RT "The structural basis of the multiple forms of human complement  
RT component C4.";  
RL Cell 36:907-914(1984).  
RN [12]  
RP SEQUENCE FROM N.A. (C4A).  
RX MEDLINE=91108039; PubMed=1988494;  
RA Yu C.Y.;  
RT "The complete exon-intron structure of a human complement component  
RT C4A gene. DNA sequences, polymorphism, and linkage to the  
RT 21-hydroxylase gene.";  
RL J. Immunol. 146:1057-1066(1991).  
RN [3]  
RP SEQUENCE FROM N.A. (C4B).  
RX TISSUE=Blood;  
RC MEDLINE=96163032; PubMed=8575831;  
RA Ulgietti D., Townsend D.C., Christiansen F.T., Dawkins R.L.,  
RA Abraham L.U.;  
RT "Complete sequence of the complement C4 gene from the HLA-A1, B8,  
RT C4A00, C4B1, DR3 haplotype.";  
RL Immunogenetics 43:250-252(1996).  
RN [4]  
RP SEQUENCE FROM N.A. (C4B).  
RX Rowen L., Dankers C., Baekin D., Faust J., Loretz C., Ahearn M.E.,  
RA Banta A., Swartzell S., Smith T.M., Spies T., Hood L.;  
RT "Sequence determination of 300 kilobases of the human class III MHC  
RT locus.";  
RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.  
RN [5]  
RP SEQUENCE FROM N.A. (C4A).  
RX Barlow K.;  
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.  
RN [6]  
RP SEQUENCE OF 1-22 AND 1056-1225 FROM N.A.  
RX MEDLINE=85156269; PubMed=3838331; Porter R.R.;  
RA Belt K.T., Yu C.Y., Carroll M.C.;  
RT "Polymorphism of human complement component C4.";  
RL Immunogenetics 21:173-180(1985).  
RN [7]  
RP SEQUENCE OF 680-756.  
RX MEDLINE=81264286; PubMed=6167582;  
RA Moon K.E., Gorzski J.P., Hugel T.E.;  
RT "Complete primary structure of human C4a anaphylatoxin.";  
RL J. Biol. Chem. 256:8685-8692(1981).  
RN [8]  
RP SEQUENCE OF 957-1044.  
RX MEDLINE=82182029; PubMed=6978711;  
RA Campbell R.D., Gagnon J., Porter R.R.;  
RT "Amino acid sequence around the thiol and reactive acyl groups of  
RT human complement component C4.";  
RL Biochem. J. 199:359-370(1981).  
RN [9]  
RP SEQUENCE OF 990-1037.  
RX MEDLINE=82150875; PubMed=6950384;  
RA Harrison R.A., Thomas M.L., Tack B.F.;  
RT "Sequence determination of the cholesterol site of the fourth  
RT component of human complement.";  
RL Proc. Natl. Acad. Sci. U.S.A. 78:7388-7392(1981).  
RN [10]  
RP SEQUENCE OF 1-21 FROM N.A.  
RX MEDLINE=94282044; PubMed=8012361;  
RA Sargent C.A., Anderson M.J., Heisler S.L., Kendall E.,  
RA Gomez-Baccobar N., Campbell R.D.;  
RT "Characterisation of the novel gene G11 lying adjacent to the  
RT complement C4A gene in the human major histocompatibility complex.";  
RL Hum. Mol. Genet. 3:481-488(1994).  
RN [11]  
RP SEQUENCE OF 1405-1431. AND SULTATION.  
RX MEDLINE=86111851; PubMed=3944109;  
RA Horton G., Sims H., Straus A.W.;  
RT "Identification of the site of sulfation of the fourth component of  
RT human complement.";  
RL J. Biol. Chem. 261:1786-1793(1986).

RN [12]  
RP STRUCTURAL BASIS OF POLYMORPHISM.  
RX MEDLINE=87080272; PubMed=2431902;  
RA Yu C.Y., Belt K.T., Giles C.M., Campbell R.D., Porter R.R.;  
RT "Structural basis of the polymorphism of human complement components  
RT C4A and C4B: gene size, reactivity and antigenicity.";  
RL EMBL J. 5:2873-2881(1986).  
RN [13]  
RP VARIANT C4A6 ALLOTYPES.  
RX MEDLINE=92242905; PubMed=1573268;  
RA Anderson M.J., Milner C.M., Cotton G.H., Campbell R.D.;  
RT "The coding sequence of the hemolytically inactive C4A6 allotype of  
RT human complement component C4 reveals that a single arginine to  
RT tryptophan substitution at beta-chain residue 458 is the likely cause  
RT of the defect.";  
RL J. Immunol. 148:2795-2802(1992).  
RN [1]  
RP FUNCTION: C4 PLAYS A CENTRAL ROLE IN THE ACTIVATION OF THE  
RN CLASSICAL PATHWAY OF THE COMPLEMENT SYSTEM. IT IS PROCESSED BY  
RN ACTIVATED C1 WHICH REMOVE FROM THE ALPHA CHAIN THE C4A  
RN ANAPHYLATOXIN.  
RN [2]  
RP FUNCTION: DERIVED FROM PROTEOLYTIC DEGRADATION OF COMPLEMENT C4,  
RN C4A ANAPHYLATOXIN IS A MEDIATOR OF LOCAL INFLAMMATORY PROCESS. IT  
RN INDUCES THE CONTRACTION OF SMOOTH MUSCLE, INCREASES VASCULAR  
RN PERMEABILITY AND CAUSES HISTAMINE RELEASE FROM MAST CELLS AND  
RN BASOPHILIC LEUKOCYTES.  
RN [3]  
RP SUBUNIT: THIS PROTEIN IS SYNTHESIZED AS A SINGLE-CHAIN PRECURSOR  
RN AND, PRIOR TO SECRETION, IS ENZYMATICALLY CLEAVED TO FORM A TRIMER  
RN OF NONIDENTICAL CHAINS (ALPHA, BETA, AND GAMMA).  
RN [4]  
RP POLYMORPHISM: HUMAN COMPLEMENT COMPONENT C4 IS POLYMORPHIC WITH AT  
RN LEAST TWO LOCI, C4A AND C4B. 13 ALLELES OF C4A AND 22 ALLELES OF  
RN C4B HAVE BEEN DETECTED. THE ALLELE SHOWN HERE IS C4A4.  
RN [5]  
RP POLYMORPHISM: THE C4A ALLELES CARRY THE BLOOD GROUP RODGERS WHILE  
RN THE C4B ALLELES CARRY THE BLOOD GROUP CHIDO.  
RN [6]  
RP DISEASE: Defects in C4A are the cause of C4a deficiency  
RN (MIM:120810).  
RN [7]  
RP DISEASE: The C4A6 allotype is totally deficient in hemolytic  
RN activity.  
RN [8]  
RP MISCELLANEOUS: C4A ALLOTYPES REACT MORE RAPIDLY WITH THE AMINO  
RN GROUP OF PEPTIDE ANTIGENS WHILE C4B ALLOTYPES REACT MORE RAPIDLY  
RN WITH THE HYDROXYL GROUP OF CARBOHYDRATE ANTIGENS.  
RN [9]  
RP MISCELLANEOUS: C4 IS A MAJOR HISTOCOMPATIBILITY COMPLEX CLASS-III  
RN PROTEIN.  
RN [10]  
RP SIMILARITY: TO C3, C5 AND ALPHA-2-MACROGLOBULIN.  
RN [11]  
RP SIMILARITY: Contains 1 anaphylatoxin-like domain.  
RN [12]  
RP This SWISS-PROT entry is copyright. It is produced through a collaboration  
RN between the Swiss Institute of Bioinformatics and the EMBL outstation -  
RN the European Bioinformatics Institute. There are no restrictions on its  
RN use by non-profit institutions as long as its content is in no way  
RN modified and this statement is not removed. Usage by and for commercial  
RN entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
RN or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
RN [13]  
RP EMBL: K02403; AAB59537.1; -  
RN EMBL: K02404; AAA59651.1; -  
RN EMBL: M59815; AAA51855.1; -  
RN EMBL: M59816; AAA51855.1; JOINED.  
RN EMBL: U24578; AAB67980.1; -  
RN EMBL: AF019413; AAB67980.1; -  
RN EMBL: AL049547; CAB89302.1; -  
RN EMBL: M14823; AAA35617.1; -  
RN EMBL: M14824; AAA52292.1; -  
RN EMBL: X77491; CAA54627.1; -  
RN PIR: I56095; C4HU.  
RN HSSP: P01031; 1KJ5.  
RN SWISS-2DPAGE: P01028; HUMAN.  
RN Genew: HGNC:1323; C4A.  
RN Genew: HGNC:1324; C4B.  
RN MIM: 120810; -  
RN MIM: 120820; -  
RN MIM: 120790; -  
RN CO: GO:000576; C:extracellular; NAS.  
RN CO: GO:0006958; P:complement activation, classical pathway; NAS.

DR GO: GO:0006954; P:inflammatory response; NAS.  
 DR GO: GO:0006937; P:regulation of muscle contraction; NAS.  
 DR InterPro: IPR002890; A2M N.  
 DR InterPro: IPR000020; Anaphylatoxin.  
 DR InterPro: IPR001599; Macroglobulin2.  
 DR InterPro: IPR001134; Netrin\_C.  
 DR Pfam: PF00207; A2M; 1.  
 DR Pfam: PF01835; A2M\_N; 1.  
 DR Pfam: PF01821; ANATO; 1.  
 DR Pfam: PF01759; NTR; 1.  
 DR ProDom: PD003264; Anaphylatoxin; 1.  
 DR SMART: SM00104; ANATO; 1.  
 DR SMART: SM00643; C345C; 1.  
 DR PROSITE: PS00477; ALPHA\_2\_MACROGLOBULIN; 1.  
 DR PROSITE: PS01177; ANAPHYLATOXIN\_1; 1.  
 DR PROSITE: PS01178; ANAPHYLATOXIN\_2; 1.  
 KW Complement pathway; Plasma; Glycoprotein; Sulfation; Signal;  
 KW Inflammatory response; Polymorphism; Disease mutation;  
 KW Blood group antigen; Thioester bond.  
 FT SIGNAL 1 19  
 FT CHAIN 20 675 COMPLEMENT C4, BETA CHAIN.  
 FT PROPEP 676 679  
 FT CHAIN 680 1446 COMPLEMENT C4, ALPHA CHAIN.  
 FT PROPEP 1447 1453  
 FT CHAIN 1454 1744 COMPLEMENT C4, GAMMA CHAIN.  
 FT PEPTIDE 680 756 C4A ANAPHYLATOXIN.  
 FT DOMAIN 702 736 ANAPHYLATOXIN-LIKE.  
 FT DISULFID 702 728 BY SIMILARITY.  
 FT DISULFID 703 735 BY SIMILARITY.  
 FT DISULFID 716 736 BY SIMILARITY.  
 FT CROSSLINK 1010 1013 Isoleucyl cysteine thioester (Cys-Gln).  
 FT MOD\_RES 1417 1417 SULFATION.  
 FT MOD\_RES 1420 1420 SULFATION.  
 FT MOD\_RES 1422 1422 SULFATION.  
 FT CARBOHYD 226 226 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 862 862 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 1328 1328 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 1391 1391 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT VARIANT 477 477 R -> W (in allotype C4A6).  
 Query Match 13.5%; Score 991.5; DB 1; Length 1744;  
 Best Local Similarity 24.7%; Pred. No. 2.8e-46;  
 Matches 409; Conservative 250; Mismatches 636; Indels 359; Gaps 65;

411 SPGSVPEVDIOONTDSSGVSPILIPQITSELQSVSAGSHPHARLTLVAAPSGCP 470  
 466 TYIQKTRDEN-TRVGSPEFL---VYSGNRKLELSMYVVSROGLVAVG---KONSTMF5 518  
 471 GFLSIERPDSPRRVGGDTLNLNRAVSGCATSEHHYIMLSRQIVPMNRPEPKTLTSVS 530  
 519 LTPENSWTPRACVIVYIIEDDGEIISDLKIPVOL-VFNKKIKLYMSKVA-EPSEKVS 576  
 531 VFDVHILAPSFYFAFYHYGD-HPVANSILRVDOAGCEGKLEISVDGAKQYRNSESVKL 589  
 577 RISVTPDLSIVGVAVDKSVNLMAAS--DITMENVHELELNTGYILGMPNMSFAVQ 634  
 590 HLE-TDGLALVALGALDITALYAAAGSKHKLNGKVFANMSYDLGGGPGSALQVQ 648  
 635 ECGI-----WYLT-----DANLTK-----DYI-----DGV- 654  
 649 AAGLAFSDGQMTLSRKRLSCPEKTKTRKRVNFOKALNEKLGQYASPTAKCCQDGV 708  
 655 -----YDAEYAEFRMEEN-----EGHIYD 674  
 709 RLPWNRSCQRAARVQPDRCBPFLSCQPAESLRKSRDKGQAGLQRALEILOEDLID 768  
 675 IHDFSLGSSPHVKHPETWTWLDTMNGYRI---YQEFVVPDSITSWATGPVISE 729  
 769 EDDIP-----VRSFPFENMLW-----RVEIVDRFQIITLWLPDSLTWETWETIGLSK 815  
 730 DLGIGLTTVEVLOAPPEFFIPLNLPSYVRGEPLLEITFNYLDATEVVKYIEKSK 769  
 816 TKGI-CVATVQVLFVEFEHLRLPLMSVRRFQLELRPVLYNLDKNLTVSIVSVEG 874  
 790 FDLIMTSSEINAGHOOTLLVPSDEGATVLPPIRPHLGEIPITVTALSP---TASDAVT 846  
 875 LCL-----AGGGGLAQVLPVAGSARPAVSVPPTAAAVSLKVAARGSEFPVGDAYS 928  
 847 QMILVKAEGIEKYSQSLIDLTDNRLOSTLKLSP-SFPPTVTGSE---RVQITAI 901  
 929 KVLQIEKEG--ALHRELVVEL--NPLDHRGRLEIPLGNSDPMMIPDGENSYVRVYASD 984  
 902 --DYLRP-----SINGLASLIRMPYGGGEOMINPAPVITYLDTLTKKO---LTDNLKEX 952  
 985 PLDTLSEGLASRQGVASLRLPRGCGEQTWITYLAPLAAASRYLDKTEQNSTLPPEYKDH 1044  
 953 ALSEMRQYORELLYQREDSFSAFGNYDPSGSTWLSAPVLRGFLEADPYIDQVNLHR 1012  
 1045 AVDLIQGVRIQGFRAQDSVYAAWLSRD--SSTWLTAFVLKXLSLAQEVGSGPEKLOE 1102  
 1013 TYTWLKGHQKSGEPMDPGVHISELQ---GNKSPVTLTAYVTSL-LGRRYKOP--- 1064  
 1103 TSNWILSQOQADGSFQPCFVLDRSMQGGIVGNDETVALTAFVTIALHGLAVFODEGAE 1162  
 1065 -----NIDVQESIHLESEFSRGISDNVTALITYALSVSQP--KAKKALMWLTWRAE 1116  
 1163 PLKQRYEASISKNSFGEESAGLGAHAAITTAALSTKAPVDLLGVAAHNNLMAAQ 1222  
 1117 QEGGMQFVWSESKLSDSMQPR-----SIDIEVAAYALSHFLQFQTSSEGI 1162  
 1223 ETGDNLYMSVYTSQSNVASEPTAPRNPDPMPQAPRLWETTRAYALL-HLL---LHEGK 1278  
 1163 PIM-----RWLSQRNLSLGGFASDTDTVALKALSEF-AALNMTERTNIQVTVGSPSS 1216  
 1279 AEMADQASAWLTRQSGFQGGFRSTQDTVALDALSAWIMSHTEEBGLAVTSL----- 1332  
 1217 PLAVVQPMANVISANGRPAICQLN-----VYVNVKASSSSRRRSI-- 1258  
 1333 -----STGRNGKSHALQJNNRQJRGLEELQFSLGSKINNVKVGNSGTLKVL 1382  
 1259 -----QNGEAPDLVAV-----KENKDLNHDV----- 1282  
 1383 TYNVLDKNTTCODLQIEVTVKGVHVEYMEANDYEDYEDLPADDDPAPLPQVTPBQ 1442  
 1283 -----NVCTSFSGP--GRSGMALMEVNLISGFNV-----PS 1311

DB 1443 LEBGRNRNRARRAPKVEQESRVAHTVTCIMRNGKVLSCMAIADVTLLSGPHALRADIE 1502  
QY 1312 EAISUSET-VKVEYDHGKLNLYLDSVNETOPCVNIPAVRNFVSTQDASVIYDYEP 1370  
DB 1503 KLTSLSDRVVSHPEHTEGPHVLLYFDSVPFTRSECVGEANQEVPPVGLVQASATLYDYNP 1562  
QY 1371 RQAVRSYNSVEYK-----LSSCDLCSVDQGCPR 1398  
DB 1563 ERRCSVFYGAPEKSRLLATLCSAEVQCAEGKCP 1596  
RESULT 11  
CO4\_MOUSE  
ID CO4\_MOUSE STANDARD; PRT; 1738 AA.  
AC P01029; Q61859;  
DT 21-JUL-1986 (Rel. 01, Created)  
DT 01-AUG-1988 (Rel. 08, Last sequence update)  
DT 15-SEP-2003 (Rel. 42, Last annotation update)  
DE Complement C4 precursor [Contains: C4A anaphylatoxin].  
GN C4.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN (1)  
RP SEQUENCE FROM N.A.  
RX MEDLINE=85298264; PubMed=3862104;  
RA Sepich D.S., Noonan D.J., Ogata R.T.;  
RT "Complete cDNA sequence of the fourth component of murine  
RT complement.";  
RL Proc. Natl. Acad. Sci. U.S.A. 82:5895-5899(1985).  
RN (2)  
RP SEQUENCE FROM N.A.  
RC STRAIN=B12.WR;  
RX MEDLINE=87309760; PubMed=3624865;  
RA Rosa P.A., Sepich D.S., Robins D.M., Ogata R.T.;  
RT "Constitutive expression of S1p genes in mouse strain B10.WR directed  
RT by C4 regulatory sequences.";  
RL J. Biol. Chem. 260:1568-1577(1987).  
RN (3)  
RP SEQUENCE FROM N.A.  
RC STRAIN=B12.WR; TISSUE=Liver;  
RX MEDLINE=89380278; PubMed=2777798;  
RA Ogata R.T., Rosa P.A., Zepf N.E.;  
RT "Sequence of the gene for murine complement component C4.";  
RL J. Biol. Chem. 264:16565-16572(1989).  
RN (4)  
RP SEQUENCE FROM N.A.  
RC STRAIN=FM; TISSUE=Liver;  
RX MEDLINE=85289294; PubMed=2993295;  
RA Nonaka M., Nakayama K., Yeul Y.D., Takahashi M.;  
RT "Complete nucleotide and derived amino acid sequences of the fourth  
RT component of mouse complement (C4). Evolutionary aspects.";  
RL J. Biol. Chem. 260:10936-10943(1985).  
RN (5)  
RP SEQUENCE OF 651-810 AND 924-1083 FROM N.A.  
RX MEDLINE=85038607; PubMed=6208559;  
RA Nonaka M., Takahashi M., Natsume-Sakai S., Nonaka M., Tanaka S.,  
RA Shimizu A., Honjo T.;  
RT "Isolation of cDNA clones specifying the fourth component of mouse  
RT complement and its isotype, sex-limited protein.";  
RL Proc. Natl. Acad. Sci. U.S.A. 81:6822-6826(1984).  
RN (6)  
RP SEQUENCE OF 1099-1142 FROM N.A.  
RX STRAIN=B10.WR, C57BL/6, C3H/He, CBA/J, B10.BR, and DBA2;  
RX MEDLINE=9053398; PubMed=2387317;  
RA Ogata R.T., Zepf N.E.;  
RT "C4 from C4-high and C4-low mouse strains have identical sequences in  
RT the region corresponding to the isotype-specific segment of human  
RL C4.";  
RL Eur. J. Immunol. 20:1607-1610(1990).  
RN (7)  
RP SEQUENCE OF 1105-1449 FROM N.A.

RX MEDLINE=85166208; PubMed=3856857;  
RA Levi-Strauss M., Toei M., Steinmetz M., Klein J., Meo T.;  
RT "Multiple duplications of complement C4 gene correlate with H-2-  
RT controlled testosterone-independent expression of its sex-limited  
RT isoform, C4-S1p.";  
RL Proc. Natl. Acad. Sci. U.S.A. 82:1746-1750(1985).  
RN (8)  
RP SEQUENCE OF 1257-1376 FROM N.A.  
RX MEDLINE=85038859; PubMed=6149581;  
RA Toei M., Levi-Strauss M., Duponchel C., Meo T.;  
RT "Sequence heterogeneity of murine complementary DNA clones related to  
RT the C4 and C4-S1p isoforms of the fourth complement component.";  
RL Philos. Trans. R. Soc. Lond., B, Biol. Sci. 306:389-394(1984).  
RN (9)  
RP SEQUENCE OF 1360-1511 FROM N.A.  
RX MEDLINE=83273751; PubMed=6192448;  
RA Ogata R.T., Shreffler D.C., Sepich D.S., Lilly S.P.;  
RT "cDNA clone spanning the alpha-gamma subunit junction in the  
RT precursor of the murine fourth complement component (C4).";  
RL Proc. Natl. Acad. Sci. U.S.A. 80:5061-5065(1983).  
RN (10)  
RP SEQUENCE OF 1-128 FROM N.A.  
RC STRAIN=FM; TISSUE=Liver;  
RX MEDLINE=86031969; PubMed=2997024;  
RA Nonaka M., Nakayama K., Yeul Y.D., Shimizu A., Takahashi M.;  
RT "Molecular cloning and characterization of complementary and genomic  
RT DNA clones for mouse C4 and S1p.";  
RL Immunol. Rev. 87:81-99(1985).  
RN (11)  
RP SEQUENCE OF 1-21 FROM N.A.  
RX MEDLINE=87017050; PubMed=3464002;  
RA Nonaka M., Kimura H., Yeul Y.D., Yokoyama S., Nakayama K.,  
RA Takahashi M.;  
RT "Identification of the 5'-flanking regulatory region responsible for  
RT the difference in transcriptional control between mouse complement C4  
RT and S1p genes.";  
RL Proc. Natl. Acad. Sci. U.S.A. 83:7883-7887(1986).  
RN (12)  
RP FUNCTION: C4 PLAYS A CENTRAL ROLE IN THE ACTIVATION OF THE  
CC CLASSICAL PATHWAY OF THE COMPLEMENT SYSTEM. IT IS PROCESSED BY  
CC ACTIVATED C1 WHICH REMOVE FROM THE ALPHA CHAIN THE C4A  
CC ANAPHYLATOXIN.  
CC  
CC -1- SUBUNIT: THIS PROTEIN IS SYNTHESIZED AS A SINGLE-CHAIN PRECURSOR  
CC AND, PRIOR TO SECRETION, IS ENZYMATICALLY CLEAVED TO FORM A TRIMER  
CC OF NONIDENTICAL CHAINS (ALPHA, BETA, AND GAMMA).  
CC  
CC -1- MISCELLANEOUS: C4 IS A MAJOR HISTOCOMPATIBILITY COMPLEX CLASS-III  
CC PROTEIN.  
CC  
CC -1- SIMILARITY: TO C3, C5 AND ALPHA-2-MACROGLOBULIN.  
CC  
CC -1- SIMILARITY: Contains 1 anaphylatoxin-like domain.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (see <http://www.ebi.ac.uk/announcements>  
CC or send an email to [license@ebi.ac.uk](mailto:license@ebi.ac.uk)).  
CC  
CC EMBL; K00019; AAA39554.1; -  
CC EMBL; M11729; AAA39506.1; -  
CC EMBL; M12968; AAA39558.1; -  
CC EMBL; M12970; AAA39555.1; -  
CC EMBL; M12972; AAA39556.1; -  
CC EMBL; M12969; AAA39559.1; -  
CC EMBL; M11789; AAA39557.1; -  
CC EMBL; K02798; AAC42021.1; -  
CC EMBL; M17440; AAA39561.1; -  
CC EMBL; M14225; AAA39563.1; -  
CC EMBL; M14226; AAA39684.1; -  
CC EMBL; X55493; CAA39112.1; -  
CC EMBL; X55495; CAA39114.1; -  
CC PIR; A24558; A24558.  
CC PIR; A29176; A29176.  
CC HSPB; P01031; IKT5.

DR MGD; MGI:88228; C4.  
 DR InterPro; IPR002890; A2M\_N.  
 DR InterPro; IPR000020; Anaphylatoxin.  
 DR InterPro; IPR001599; Macroglobulin2.  
 DR InterPro; IPR001134; Nectrin\_C.  
 DR Pfam; PF00207; A2M; 1.  
 DR Pfam; PF01835; A2M\_N; 1.  
 DR Pfam; PF01821; ANATO; 1.  
 DR Pfam; PF01759; NTR; 1.  
 DR ProDom; PD003264; Anaphylatoxin; 1.  
 DR SMART; SM00643; ANATO; 1.  
 DR PROSITE; PS00477; ALPHA\_2\_MACROGLOBULIN; 1.  
 DR PROSITE; PS01177; ANAPHYLATOXIN\_1; 1.  
 DR PROSITE; PS01178; ANAPHYLATOXIN\_2; 1.  
 KW Complement pathway; Plasma; Glycoprotein; Sulfation; Signal;  
 KW Inflammatory response; Thioester bond.  
 FT SIGNAL 1 19  
 FT CHAIN 20 673 COMPLEMENT C4, BETA CHAIN.  
 FT PROPEP 674 677 COMPLEMENT C4, ALPHA CHAIN.  
 FT CHAIN 678 1443  
 FT PROPEP 1444 1447  
 FT CHAIN 1448 1738  
 FT PEPTIDE 678 753  
 FT DOMAIN 700 734  
 FT DISULFID 700 726  
 FT DISULFID 701 733  
 FT DISULFID 714 734  
 FT CROSSLINK 1006 1009  
 FT MOD\_RES 1413 1413  
 FT MOD\_RES 1416 1416  
 FT MOD\_RES 1417 1417  
 FT CARBOHYD 224 224  
 FT CARBOHYD 743 743  
 FT CARBOHYD 1387 1387  
 FT CONFLICT 132 132  
 FT CONFLICT 327 327  
 FT CONFLICT 570 570  
 FT CONFLICT 720 720  
 FT CONFLICT 739 740  
 FT CONFLICT 838 838  
 FT CONFLICT 993 993  
 FT CONFLICT 1043 1043  
 FT CONFLICT 1119 1119  
 FT CONFLICT 1190 1190  
 FT CONFLICT 1324 1324  
 FT CONFLICT 1401 1401  
 FT CONFLICT 1442 1442  
 FT CONFLICT 1453 1453  
 SQ SEQUENCE 1738 AA; 192870 MW; D1E02AE7AB42BFF CRC64;  
 Query Match 13.4%; Score 987; DB 1; Length 1738;  
 Best Local Similarity 24.9%; Pred. No. 4.8e-46;  
 Matches 415; Conservative 245; Mismatches 643; Indels 366; Gaps 64;  
 QY 10 AHLLCVCAALAVAPGPRFLVATGIRPGNVTIGVLEHPSQVTVKAE----- 62  
 DB 8 AMVVSFCASSL---QKPRLLLFSPSVNLGTPLVGVQLDAPRGQ-EVKSQVFLRNPKG 63  
 QY 63 -----LKTASNLTVSVLEA--EGVFEKSGFKTLTLPSPRLNSADEIYELRVGT 110  
 DB 64 GSGSPKDFKLSGDDPVLISLEVPLEDRVSCGLFDRBRAHIDLVAGSP--WLRNTAFK 121  
 QY 111 QDELLFNSNSTRLSSETKRIISVFIOTDKALYKPKQEVKFRIVTLFSDFKPYKTSNLIIKD 170  
 DB 122 ATE---TQGVNVLTFSSRRGHIFVQTDQPIYNPGQVRVRYFALDQKMRPSTDFLTIVEN 178  
 QY 171 PKS-NLIQOMLSQSDGIVSKTQSLSHPTLDGMSIOGVND---QTYVGSFOVSEVYL 226  
 DB 179 SHGRARLVKKEIFITSTISFDQFTIPDISEP--GTWKLSARPSDGLSENRSTHEFVKKIVL 236  
 QY 227 PKFEVTLQTP-----LYCSMNSKHLNGITITAKTYTGKPVKGDVLTLEPLSFMGKKNI- 280

DB 237 PNEFVKI-TPWKPYIILVPSNSDEIQLDIQARIYTGKPVGVAVYTRPALMDEGKTFELR 295  
 QY 281 ---TKTEKINGSANFSNDEEMGVNDSSN-----GLSEYLDLSPGVEILTITVTES 330  
 DB 296 GLETOAKLVGRGRHISISKDQFALDKINIGRDEGLRLY-----AATAVIES 345  
 QY 331 VTGISRNVS--TNVFKQHDYIIEFFDYTVLKESLNFATATVKYTRADGNQLTLEERNV 389  
 DB 346 PGEMERAEELTSWRFPVSASFSLDSRTKRLVPGAHFLDALQVEMSGSE-----ASN 399  
 QY 390 VITVTOQNTYEVSGNSGNQKNEANQKINYTPVQSGTFKIEEPILEDSELOKAYFLG 449  
 DB 400 PVKVS---ATLVSGDS--QVLDIQSTT---GIGVSISEPIPTVELLVSAGS 449  
 QY 450 SKSMVAHSLFKSPSK--TYIOLKTRD-ENIKYGSPEL-----VSGNKKLSELYVVS 502  
 DB 450 LYPALRLTYQAPPSKGTGLSIEPDRPSVDGFILNLQVGVGAPAPFPHHYMTIS 509  
 QY 503 RGQLVAVGKO---NSTWFSLTTPNSWTPKACVIVYIIEDDGELISDVLKIPVQ---LVFK 556  
 DB 510 RGQIMAMGREPRKTVTSVSLVDHQLABSFYFVAFYH--QGHFVANSLLINISRDCEG 568  
 QY 557 NKIKLWYSKYKAPSEKVSRLISVTPQDSIVGIVAVDKSVNLM--NANSDITENYVHEL 614  
 DB 569 LQKLVGAK-EYRNADMKRLIQ-TDSKALVALGAVDTALVAVGSRSHKPLDMSKVEVI 626  
 QY 615 ELVNTGYLLFMFNSFAVPOEGLMWLTDLANTKDYID----- 652  
 DB 627 NSTNVGCGPGGDALQVFDAGLAFSDGDRLTQTRBEDLSCPKEKKSROKRVNFOKAVS 666  
 QY 653 ---GVYDN-----AAYERFMEZ 667  
 DB 687 EKLGGVSSPPARKCCDQMKTLPMKRTCEQRAARVPOACREFFLSCCKRABDLRRNQTR 746  
 QY 668 NEGH-----YDIDHPSLSSSPHVKHPEWTW---LDITMGRIYQEF 710  
 DB 747 SQHLARNNNMLQEBDLIDEDDIL-----VRTSPENMLMVEPDS-----KLLT 794  
 QY 711 VTPDSTSVVATGPIYSBDLGLGLTTPELQAFQFFIFLNLPSVIRGEFEALFTI 770  
 DB 795 VMLPDSWTTWEIHGVSSLSKSGU-CVAKPTRVAVFRFHMLPLSIRRFQEBELRPVL 853  
 QY 771 FNYLKATEVYVIEKSDKEDILMTSEINATGHQOQLLVPSHDGATVLPPIRPHLGEI 830  
 DB 854 YNVLND--DVAVSVHTPVSGGLAGGMYA-----QGVTVPASARPAVSVPVTAANY 907  
 QY 831 PITVTLASP-TASDAVYQMLVKAEGIEKYSQSIILDLTD-NRLOSTLKLTSFSPFN- 887  
 DB 908 PLKVVARGVFDLDGAVSKILQIEKEG--AIHRELVYNLDPLNNLGRITLE-IPGSSDPNI 964  
 QY 888 -----TYMSEVQVLTALGDVLPISINGLSLIRMPVCGCGGOMNMFANPIYILD 937  
 DB 965 VPDGPRSSLVRAVASEPLETMSSEGAISP--GGVASLRLRPQCCAQOTMYLAPTLASN 1022  
 QY 938 YLTKKKO---LITNLLEKALSPKOCYQRELLVQREDSGSARANYDPGSGTWLSAFVLR 994  
 DB 1023 YLDRTEQMSKLSFETDHDVNDLQKGYMRLQCRKKNDSGAMLHND--STWLTAFVLK 1080  
 QY 995 CLEADPYIDIDONVLRHTYMLKGHOKSNGEFPWDPGRVHSELQG--GNKSPVTLVAY 1051  
 DB 1081 ILSLAQEQVNSPEKLOETASWMLAQQLGDSFHDPCPVLRAMQGLVGSDETVALTAF 1140  
 QY 1052 IYVSL-----LGRKYQPNIDVQESIHFLSESPSRGSDNYTIALITYALSSVG 1100  
 DB 1141 VVIALHGHGLVDODDAKQKNNVEASITKANSFLQOKASAGLLGAHAAAITVALTLT- 1199  
 QY 1101 SPKAKAL-----NMLTWRADEGGMQFW---VSSESK-LSDSWQPRS-----LD 1141  
 DB 1200 --KASEDLRVANINSLMAAEFGEHLVNGVLVGSQDKVLRRTARSRPEVPQAPALW 1257  
 QY 1142 IEVAAVALSHFLQFQTSBGPIM-----RMLSGRQNSLGGFASSTQDTYVALKALSEF-A 1195



Db 1258 IETTAVALLHLIR-----EKGKMDAKASWILTHQSGFHGAFRRSTODTVTLIDALSANI 1313  
 Qy 1196 ALMNTERTNIQVT-----VTGSSPSPLAVQVPAVNISANGFCF 1235  
 Db 1314 ASHTTEBKAKYTLTSMGNGKLTGHILNNHNVQVGLBEELKFSLSGTSISVVEGNSKOT 1373  
 Qy 1236 -----AICQLNVVNVYKASGS----- 1251  
 Db 1374 LKILRTVNVLDKMKNTTCC-DLQIEVKVGTGAVEYAMANDNEYEDYDMPAADPSVPLQPV 1432  
 Qy 1252 -----SRRRSIOQNEAFDDIVANKENDLNVHDLNVCTISFSGP-GRSGMALME 1300  
 Db 1433 TPLQJFEGRRSRRRR-----APKVAEBOESRQVYTCIWRNGKGLSGMAIAD 1481  
 Qy 1301 VNLISGFNV-----PSEAISET-VKVEYEDGKNTLYLDSVNETQFCVNIPIAVNFKVS 1355  
 Db 1482 ITLISGFTALRADLEKTLTSLSDRYVSHFPTDPRHVLIVDSPTTRECVCFGASQGBVVG 1541  
 Qy 1356 NTQDASVSIVDYEPERRQVRSYNSVVK-----LSSCDLCSDVQGCPR 1398  
 Db 1542 LVQPSASVAVDYXSPDHKCSFYAAPTKSQLATLCSGVCVCCAGSKCP 1590  
 RESULT 12  
 CO3\_ONCMY STANDARD; PRT; 1640 AA.  
 AC P98093;  
 DT 01-FEB-1996 (Rel. 33, Created)  
 DT 01-FEB-1996 (Rel. 42, Last annotation update)  
 DT 15-SEP-2003 (Rel. 33, Last sequence update)  
 DE Complement C3-1 (Contains: C3A anaphylatoxin) (Fragment).  
 OS Oncorhynchus mykiss (Rainbow trout) (Salmo gairdneri).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;  
 OC Proactinopterygii; Salmoniformes; Salmonidae; Oncorhynchus.  
 NC NCBI\_TaxID=8022;  
 RN [1]  
 RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.  
 RC TISSUE=Liver;  
 RX MEDLINE=94065166; PubMed=8245455;  
 RA Lambiris J.D., Lao Z., Pang J., Alsenz J.;  
 RT "Third component of trout complement. cDNA cloning and conservation  
 of functional sites.";  
 RL J. Immunol. 151:6123-6134(1993).  
 CC -1- FUNCTION: C3 PLAYS A CENTRAL ROLE IN THE ACTIVATION OF THE  
 COMPLEMENT SYSTEM. ITS PROCESSING BY C3 CONVERTASE IS THE CENTRAL  
 REACTION IN BOTH CLASSICAL AND ALTERNATIVE COMPLEMENT PATHWAYS.  
 CC AFTER ACTIVATION C3B CAN BIND COVALENTLY, VIA ITS REACTIVE  
 THIOLESTER, TO CELL SURFACE CARBOHYDRATES OR IMMUNE AGGREGATES.  
 CC -1- SUBUNIT: C3 PRECURSOR IS FIRST PROCESSED BY THE REMOVAL OF 4 ARG  
 RESIDUES, FORMING TWO CHAINS, BETA & ALPHA. LINKED BY A DISULFIDE  
 BOND. C3 CONVERTASE ACTIVATES C3 BY CLEAVING THE ALPHA CHAIN,  
 RELEASING C3A ANAPHYLATOXIN & GENERATING C3B (BETA CHAIN + ALPHA'  
 CHAIN).  
 CC -1- SIMILARITY: TO C4, C5 AND ALPHA-2-MACROGLOBULIN.  
 CC -1- SIMILARITY: Contains 1 anaphylatoxin-like domain.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation-  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC -----  
 CC EMBL; L24433; AAB05029.1; ALT\_INIT.  
 DR PIR; I51339; I51339.  
 DR HSSP; P01024; ICD3.  
 DR InterPro; IPR002890; A2M N.  
 DR InterPro; IPR000020; Anaphylatoxin.  
 DR InterPro; IPR001599; MacroglobinA2.  
 DR InterPro; IPR001134; Netrin\_C.  
 DR Pfam; PF00207; A2M; 1.

DR Pfam; PF01835; A2M N; 1.  
 DR Pfam; PF01821; ANATO; 1.  
 DR Pfam; PF01759; NTR; 1.  
 DR ProDom; PD003264; Anaphylatoxin; 1.  
 DR SMART; SM00104; ANATO; 1.  
 DR SMART; SM00643; C345C; 1.  
 DR PROSITE; PS00477; ALPHA 2 MACROGLOBULIN; 1.  
 DR PROSITE; PS01177; ANAPHYLATOXIN 1; 1.  
 DR PROSITE; PS01178; ANAPHYLATOXIN 2; 1.  
 KM Complement pathway; Complement alternate pathway; Plasma;  
 KM Inflammatory response; Glycoprotein; Thioester bond.  
 FT NON\_TER 1 1640  
 FT CHAIN 1 1640  
 FT CHAIN 1 642  
 FT CHAIN 647 1640  
 FT PEPTIDE 647 722  
 FT CHAIN 723 1640  
 FT PEPTIDE 723 931  
 FT PEPTIDE 932 1278  
 FT PEPTIDE 932 1033  
 FT PEPTIDE 1034 1278  
 FT PEPTIDE 1279 1295  
 FT SITE 722 723  
 FT SITE 931 932  
 FT SITE 1278 1279  
 FT SITE 1295 1296  
 FT DOMAIN 668 703  
 FT DOMAIN 1420 1430  
 FT DISULFID 536 797  
 FT DISULFID 603 638  
 FT DISULFID 669 702  
 FT DISULFID 682 703  
 FT DISULFID 853 1488  
 FT DISULFID 1079 1135  
 FT DISULFID 1335 1464  
 FT DISULFID 1481 1486  
 FT DISULFID 1493 1563  
 FT DISULFID 1510 1638  
 FT DISULFID 1614 1623  
 FT CARBOHYD 164 164  
 FT CROSSLINK 988 991  
 SQ SEQUENCE 1640 AA; 162104 MW; 0965B4FAFB87812 CRC64;  
 Query Match 13.4%; Score 984.5; DB 1; Length 1640;  
 Best Local Similarity 24.1%; Pred. No. 6.1e-46;  
 Matches 387; Conservative 298; Mismatches 571; Indels 349; Gaps 67;  
 Qy 18 AALVAPPRRLVAPAGIIRRGAVNTGVLELHCPGQVTVKAKELTKASNLTVSVLAE 77  
 Db 1 AALQV-----LSAPNLLRVGSNENIFVESQDHVGGLNKL-VNKNHPQOSKELASKS 52  
 Qy 78 GVPEK-GSPKTLT-----LPSLPLNSADEIELRYRTGTODEILFSSNSTRLSFETKR 128  
 Db 53 VVLQGANNFQMTQVLVIGRGLVDDPKQKQVVLQAGRPDDLLEKVV-----VSFGSGY 107  
 Qy 129 ISVFIQDKALYKPKQEVKPRIVTLFSDPKY-----KTSINILIKDPKSN 174  
 Db 108 I--FIQDKKITVTASTLVHYVVFMTPLBEFLTREIFEDQEVAKNKEIAYVVEIMTPENI 165  
 Qy 175 LIQQLMSQSGSLDGLVSKTFOSSHPIIGDMSIQVQV---DQTVYQSGVQSEVYLKPREV 231  
 Db 166 TIFRRI-VNPKGVKSGQFKLPDIVSFTVHWIVRFPSTPQKTSSESEVEVEYVLPSEV 224  
 Qy 232 TLQTP--LYCSMNSKHLNGITTAKYTGKPVKGDVTLTF-----LPLSPFWG--- 275  
 Db 225 SL-TPAKAFVYVDDNDLVDITARIYLVKEVYTGIGYVVFVITTBSEKKSPPASIGRBEI 283  
 Qy 276 -----KKKNTTKTF-KINGSANPSEFNDEKKNVMDSSNGLSSEYLDSSPGQVEILT 325  
 Db 284 KDQGVACLAKKEHITQTFPKIHDLVKOSI-----FVSVSVLTEGGGEVWEAEKRG-IQIVT 338

Qy	326	TVTSEVTG1SRNVSTNVFQKHDIYIIIEFDYTTVLKPSLNFPAKYATKVRADQNGQLTBER	385
Db	339	S-----PYILKRPRPKYKPCGMPVUSYITINP-----	368
Qy	386	RNNVVITYTORNYTEYWGSGNSGNOKMEAVOKINYTVPOSGTFKIEPILDESELQKA	445
Db	369	-NSPAIVEVEVTPHAKGVTRAN-----GPAKIPMTVASATELVTIV	411
Qy	446	-----YFLGSSSMAVHSLFSPKTIYQLKTRDENIKVSPPELVVS-GNKL--	493
Db	412	KTKPDGDRPOQTGGGATMALP--YRTSTKNFLHVDNSNELKIDPICKIDMLGPTTIPN	469
Qy	494	KELSYNVSRGOLVAVG---KONSTMFSLT--PENSMTPKACVIVYIYIEDGEIISDVLK	548
Db	470	HDLTTFMFSRGOLVAVGFRKQGNALVTLSPVSKELLPSFRTIYAVHYGAADLVADSVW	529
Qy	549	IPVOLFKNKIKLYMSKVKA--EPSEKVSRLISVTPQPSIVIGIYAVDKSVMLMNASNDIT	606
Db	530	VDIKVSCMGSLKVTSTRPKASYEPFARASLTIT--GDPGAKVGLVAVDKGVYVLYNSKHRLT	588
Qy	607	MENVVHELELYTGYLGMFANISFAVPOECGL-----WUTLDAN-----	645
Db	589	QTKIWDITIEKHDTGCTAGGADNMGVFADAGLVPEFNTAKGTGIRTDSCPVSSRRRAV	648
Qy	646	-----LTKD-YIDGYDN-----AEXAE	662
Db	649	TISDVIYTMASKYHGLAKECCVDGGRDNTMGYTCDRRAQYISDGVCVQAFVLCCTEMAS	708
Qy	663	RFMEBNEGHIV-----DIDFSLGSSPHV-RKHPETWILDTNMGYRIYQ--EPEVT	712
Db	709	KKIESKQOALLNSREEDDDAVYRSRSDIYRSQFPESMMEDINLEBCRQNGHCST	768
Qy	713	-----VPDSITSVATGATVISEDGLGLTTPEYLOAFQDFEFLMLPVSIRNGEPA	765
Db	769	SVIRNNFLKDSITTTQITAIASLSTKHGI-CVADPEMIVLKEFFIDLKLPASVANEQLE	827
Qy	766	LEITFNTLKATEVKYIIEKSDKFDILMTSEINATHOOTLVPSRDAQTVLEPIRPT	825
Db	828	VKALIHNSSEDPITYKVELMENG-----VCSASKKGYKROCVNMDPMSTRVYPVITIPM	883
Qy	826	HLG-ELIPTVYALSPASDAVQOMILYKABEIEKYSQSILDLTDRLQSTLTKLSFS	883
Db	884	KLGLHSIEVKA.SVKNKSGNDGKYKRLRVABEVVLKKEITVNLNVKHKGEPT-SHISG	942
Qy	884	FPNPTVTGSR--VOITAIQD---VLGPSING--LASLIRMPYGCBBQNMINFANITY	934
Db	943	VPRNQVPNSDADTLISVTA-GEQTSVLVEQAISGDSLIAVQVPGCEQOMIYMTLEPVI	1001
Qy	935	ILDYITKKQOLD---NLKEXALSTMROGYQRELLYQREDSFFAFNGYVDPGSGTWLSAF	991
Db	1002	ATHYIDNTKKMEDIGLDRKNTAIKIKINIGYQORQLARREKDSYAAW--VSRQSTWMLTAY	1055
Qy	992	VLRCFLIADPYIDIDONYLHRTYTWL-KQHOKSNGEPMFDPGRVHSELQG--GNKSPVT	1047
Db	1060	VVKYFAMSTLISQENVLCYAKVLLINTQOPDITFHEFAPVHIAEMTGNVRSQDNDA5	1119
Qy	1048	LTAYI-----VTSLLGYKRYQPNIDVQESIHFLSEFSRGISDNTYTLAIT	1093
Db	1120	MTAFVLIMQEASSVCEQSVNSLPG-----SMAKAVAYLEKRLPH-LTNPVAVAMTS	1170
Qy	1094	YALSSVSGPKAKKALNMTLWRAEO--EGGMQFVWSSSKLSDSQPSLDIEVAAVYLL	1155
Db	1171	YALNAGLKNKETLLKFASPOLDHPWPBGQYQ-----TLEATSYALL	1213
Qy	1151	SHFLQFQTSSEGIPIRWLTSRONSLJGFSAPQDTPVALKALSEFALNMNTER--TNIQ	1207
Db	1214	ALVYKAFEEKGPIYKMLNKQKKGCGGSGTOSTITMVFQAAVEYSHYKDLKDFPLNINL	1273
Qy	1208	TVTGPSBSPSLAVVQPMAVNISANGFGALCOLNV---YNYKASGSRRRRSI-----	1258
Db	1274	EVAGRAS-----VTKMSIN-NKQNGFHRTDKVNSIDDLQVTKASGNGEATLISVTLTYVA	1328
Qy	1259	-----ONQEAFLDLDVAVKENKDINHVD-----LNVCTSFSGFGRSG-MALNEVNILS	1305

```

Db      1327 LPEKSDSCBESFDLSVLTIK-MDRTSHEDAKESFMILITVULYKNSSEDDATMSIDILTLT 1365
Qy      1306 GFMWPSERIS-LSF-----TYKKVEYD-----HGKLNLYDSVN--ETQFCVNIPAVRNF 1352
Db      1386 GFIVDTDLNLNLSKGRERYIEKPEMDKVLSEKSGSLIYLDKVSHEKEDRISFKIHVQ-- 1443
Qy      1353 KVSNTQDASVSIVDYEPFRQAVNSYSE-----VLTSSCDLCS 1391
Db      1444 EGVGLQPAVASVVEYNOKR-CVKEFYHPQREBGTLISRLCLGDVCT 1487

RESULT 13
CO3_NAJNA
ID CO3 NAJNA STANDARD; PRT; 1651.AA.
AC Q01833;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Complement C3 precursor [contains: C3a anaphylatoxin].
C3.
OS Naja naja (Indian cobra).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Lepidoptera; Squamata; Sclerozoidea; Serpentes; Colubroidea;
OC Elapidae; Elapinae; Naja.
NCBI_TaxID=35670;
(1)
SEQUENCE FROM N.A.
RP TISSUE=Liver;
RX MEDLINE=93056528; PubMed=1431125;
RA Fritzlinger D.C., Connolly M., Petrella E.C., Bedehorst R.,
RA Vogel C.W.;
RT "Primary structure of cobra complement component C3.";
RL J. Immunol. 149:3554-3562(1992).
CC -1- FUNCTION: C3 PLAYS A CENTRAL ROLE IN THE ACTIVATION OF THE
CC COMPLEMENT SYSTEM. ITS PROCESSING BY C3 CONVERTASE IS THE CENTRAL
CC REACTION IN BOTH CLASSICAL AND ALTERNATIVE COMPLEMENT PATHWAYS.
CC AFTER ACTIVATION C3B CAN BIND COVALENTLY, VIA ITS REACTIVE
CC THIOLESTER, TO CELL SURFACE CARBOHYDRATES OR IMMUNE AGGREGATES.
CC -1- FUNCTION: DERIVED FROM PROTEOLYTIC DEGRADATION OF COMPLEMENT C3,
CC C3A ANAPHYLATOXIN IS A MEDIATOR OF LOCAL INFLAMMATORY PROCESS. IT
CC INDUCES THE CONTRACTION OF SMOOTH MUSCLE, INCREASES VASCULAR
CC PERMEABILITY AND CAUSES HISTAMINE RELEASE FROM MAST CELLS AND
CC BASOPHILIC LEUKOCYTES.
CC -1- SUBUNIT: C3 precursor is first processed by the removal of 4 Arg
CC residues, forming two chains, beta and alpha, linked by a
CC disulfide bond. C3 convertase activates C3 by cleaving the alpha
CC chain, releasing C3a anaphylatoxin and generating C3b (beta chain
CC + alpha' chain) (By similarity).
CC -1- SIMILARITY: TO C4, C5 AND ALPHA-2-MACROGLOBULIN.
CC -1- SIMILARITY: Contains 1 anaphylatoxin-like domain.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
DR EMBL, L02365; AAA49385.1; -.
DR HSSP, P01024; IC3D.
DR InterPro, IPR002890; A2M_N.
DR InterPro, IPR000020; Anaphylatoxin.
DR InterPro, IPR001840; Anaphylatoxin.
DR InterPro, IPR001599; MacroglobulinA2.
DR InterPro, IPR001134; Netrin_C.
DR Pfam, PF00207; A2M_1.
DR Pfam, PF01835; A2M_N; 1.
DR Pfam, PF01821; ANATO; 1.
DR Pfam, PF01759; NTR; 1.
DR PRINTS, PRO00004; ANAPHYLATOXN.
DR PRODom, PD003264; Anaphylatoxin; 1.

```



RESULT 14  
 CO3\_RAT STANDARD, PRT: 1663 AA.  
 ID CO3\_RAT  
 AC P01026;  
 DT 21-JUL-1986 (Rel. 01, Created)  
 DT 01-AUG-1990 (Rel. 15, Last sequence update)  
 DT 15-SEP-2003 (Rel. 42, Last annotation update)  
 DE Complement C3 precursor [Contains: C3a anaphylatoxin].  
 GN C3.  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 OX NCBI\_TaxID=10116;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Wistar; TISSUE=Liver;  
 RX MEDLINE=90245672; PubMed=2336397;  
 RA Misumi Y., Sohma M., Ikehara Y.;  
 RT "Nucleotide and deduced amino acid sequence of rat complement C3,"  
 RL Nucleic Acids Res. 18:2178-2178(1990).  
 RN [2]  
 RP SEQUENCE OF 671-748.  
 RX MEDLINE=79062262; PubMed=309768;  
 RA Jacobs J.W., Rubin J.S., Hugel T.E., Bogardt R.A., Mariz I.K.,  
 RA Daniels J.S., Daughaday W.H., Bradshaw R.A.;  
 RT "Purification, characterization, and amino acid sequence of rat  
 anaphylatoxin (C3a)."  
 RL Biochemistry 17:5031-5038(1978).  
 RN [3]  
 RP SEQUENCE OF 1316-1595 FROM N.A.  
 RX MEDLINE=89380332; PubMed=2674144;  
 RA Sundstrom S.A., Komm B.S., Ponce-De-Leon H., Yi Z., Teuscher C.,  
 RA Lytle C.R.;  
 RT "Retrogen regulation of tissue-specific expression of complement C3,"  
 RL J. Biol. Chem. 264:16941-16947(1989).  
 CC -1- FUNCTION: C3 PLAYS A CENTRAL ROLE IN THE ACTIVATION OF THE  
 COMPLEMENT SYSTEM. ITS PROCESSING BY C3 CONVERTASE IS THE CENTRAL  
 REACTION IN BOTH CLASSICAL AND ALTERNATIVE COMPLEMENT PATHWAYS.  
 CC AFTER ACTIVATION C3B CAN BIND COVALENTLY, VIA ITS AGGREGATES,  
 CC THIOLESTER, TO CELL SURFACE CARBOHYDRATES OR IMMUNE AGGREGATES.  
 CC -1- FUNCTION: DERIVED FROM PROTEOLYTIC DEGRADATION OF COMPLEMENT C3,  
 CC C3 ANAPHYLATOXIN IS A MEDIATOR OF LOCAL INFLAMMATORY PROCESS. IT  
 CC INDUCES THE CONTRACTION OF SMOOTH MUSCLE, INCREASES VASCULAR  
 CC PERMEABILITY AND CAUSES HISTAMINE RELEASE FROM MAST CELLS AND  
 CC BASOPHILIC LEUKOCYTES.  
 CC -1- SUBUNIT: C3 PRECURSOR IS FIRST PROCESSED BY THE REMOVAL OF 4 ARG  
 CC RESIDUES, FORMING TWO CHAINS, BETA & ALPHA, LINKED BY A DISULFIDE  
 CC BOND. C3 CONVERTASE ACTIVATES C3 BY CLEAVING THE ALPHA CHAIN,  
 CC RELEASING C3A ANAPHYLATOXIN & GENERATING C3B (BETA CHAIN + ALPHA  
 CC CHAIN).  
 CC -1- SIMILARITY: TO C4, C5 AND ALPHA-2-MACROGLOBULIN.  
 CC -1- SIMILARITY: Contains 1 anaphylatoxin-like domain.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.isb.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL, X52477; CA36716.1; -;  
 DR EMBL, M29866; AAA40837.1; ALT\_SEQ.  
 DR PIR, S15764; C3RT.  
 DR PDB, 1QOF; 31-JUL-00.  
 DR PDB, 1QSF; 31-JUL-00.  
 DR InterPro; IPR002890; A2M\_N.  
 DR InterPro; IPR000020; Anaphylatoxin.  
 DR InterPro; IPR001599; Macroglobulin2.  
 DR InterPro; IPR001134; Netrin\_C.  
 DR Pfam; PF00207; A2M; 1.

DR Pfam; PF01835; A2M\_N; 1.  
 DR Pfam; PF01821; ANATO; 1.  
 DR Pfam; PF01759; NTR; 1.  
 DR ProDom; PD003264; Anaphylatoxin; 1.  
 DR SMART; SM00104; ANATO; 1.  
 DR SMART; SM00643; C345C; 1.  
 DR PROSITE; PS00477; ALPHA\_2\_MACROGLOBULIN; 1.  
 DR PROSITE; PS01177; ANAPHYLATOXIN\_1; 1.  
 DR PROSITE; PS01178; ANAPHYLATOXIN\_2; 1.  
 KW Complement pathway; Complement alternate pathway; Plasma;  
 KW Inflammatory response; Glycoprotein; Signal; 3D-structure;  
 KW Thioester bond.  
 FT SIGNAL 1 24  
 FT CHAIN 25 1663 COMPLEMENT C3.  
 FT CHAIN 25 666 BETA CHAIN.  
 FT CHAIN 671 1663 ALPHA CHAIN.  
 FT PEPTIDE 671 748 C3A ANAPHYLATOXIN.  
 FT CHAIN 749 1663 C3B (ALPHA' CHAIN).  
 FT SITE 748 749 CLEAVAGE (BY C3 CONVERTASE).  
 FT DOMAIN 693 728 ANAPHYLATOXIN-LIKE.  
 FT DISULFID 558 816 INTERCHAIN (BY SIMILARITY).  
 FT DISULFID 626 661 BY SIMILARITY.  
 FT DISULFID 693 720 BY SIMILARITY.  
 FT DISULFID 694 727 BY SIMILARITY.  
 FT DISULFID 707 728 BY SIMILARITY.  
 FT DISULFID 873 1513 BY SIMILARITY.  
 FT DISULFID 1101 1158 BY SIMILARITY.  
 FT DISULFID 1358 1489 BY SIMILARITY.  
 FT DISULFID 1389 1458 BY SIMILARITY.  
 FT DISULFID 1506 1511 BY SIMILARITY.  
 FT DISULFID 1518 1590 BY SIMILARITY.  
 FT DISULFID 1537 1661 BY SIMILARITY.  
 FT CROSSLINK 1010 1013 Iso-glutamy1 cysteine thioester (Cys-Gln).  
 FT CARBOHYD 939 939 N-LINKED (GLCNAC. . .) (PROBABLY).  
 FT CARBOHYD 1617 1617 LK -> KL (IN REF. 2).  
 FT CONFLICT 721 722  
 SQ SEQUENCE 1663 AA; 186460 MW; 2897CCB143CDD4BC CRC64;  
 Query Match 13.1%; Score 966; DB 1; Length 1663;  
 Best Local Similarity 24.3%; Pred. No. 6.4e-45; Index 326; Gaps 68;  
 Matches 391; Conservative 284; Mismatches 609;  
 QY 3 GPPLTAHMLLCVCTALAAVAPGRPLVTAAGIIRRGNTVIGVELLE--HCPGQTVK 59  
 DB 6 GSQLV--LLLLASLALGSPMVSITPNTVRLSESEFTFLHMDAGDVPVTVQ 62  
 QY 60 AELIKTASNLTVLVE--AEVGFEKSGFKTLTLPSPLNADLEYELRVTRTODELIFSN 118  
 DB 63 DFLKKQVLTSKTVLTGATGHLNRVPIK--IPASKEFNADKHKVTVVANFGATVEK 119  
 QY 119 STLSPEPKRISVFIDPDKALYKPKOEKFRIVTLFSDPKPYKTSINILHKDPKSNLIQ 178  
 DB 120 AVLVSPQSGYL--FIQTDKITYPGSTVFRIFTVDNNLLPVKRTIVVILETDGVPDKR 177  
 QY 179 --WLSQSDLDLVISCTQLSSHPILGDSIQV--QVNDQTYGSPQVSEVYLPKFEVTLQ 234  
 DB 178 DILSSHNQVGIPLSNINPELVMGQKIRAFEPHAKQFFSAFVEKVEVLVSFEVLVE 237  
 QY 235 --TPLYCSNMSKILNGITTAKYKPVKGDVTLTF-----LPLSGFKKKNIITK 282  
 DB 238 PTEKFYIIHPPKLEVISITARFLYGRKVVDTAFVIGVODEDKISILAS-----LTR 290  
 QY 283 TPKINGSANFSFDEMKVNDSSNGLSSEYLDSSPGPVELITLVESVYGISRNVSTNV 342  
 DB 291 VLIEDSGEAVLS--RKVLMDG-----VRPSP-----EALGKSLVYSVTV 330  
 QY 343 FFKQ-----HDYIIFFDYTVLKEPSLNETATVATVTRADGNOLTEERR 386  
 DB 331 IILHSGDMVAERSGIPYIVSPYQIHFTKPKFKFXAMPDLDVAVFTNPQGS----PAR 386  
 QY 387 NNVTIVTQNNTEYWSGNS--GNQKREAVOKINTVPOS-----GTFKIEPFLIED 437  
 DB 387 ---VPVTVQ-----GSDAQALQTDDGVAKLISVNTPNRQPLITIVSTKKEGIPDARQ 435

```

QY 438 SGE-LQKAYPLGSSMAVHSLFKSPSTYIQLKTRDENIVGSPFELVSGNRLKEL 496
DB 436 ATRTQAOQPYSTHMSNNYIHL-----SVSRVELKPGD-NLVNFIHLRTDAGQEKIRYX 489
QY 497 SYMVTSRQGLVAVK-----ONSTMFSL--TPENSWTPKACVIYV--IEDDG--EIS 544
DB 490 TYLVNKKGLLAKAGQVREPPGQDLVLSLPIPE--FISFRLVAYTLIGANGQREVA 547
QY 545 DYLKIPVQ-----LVFKKIKLYMSKVYAESEKSLRISTOPDSIYGIYAVVSVL 598
DB 548 DSWMTVDVDSVCGTILVVKDPR---DNQPAHGHTLLIENQ--GARGLVAVDKGVV 603
QY 599 MNASDITMENVHLELYNTG-----YILGEMNSPAVQ-----ECG 637
DB 604 LNKKKKLQSKIMDVEKADICTPGSKGVYAGVFMDAGLTFKTNQGLQDQREDECAK 663
QY 638 -----LWLTDLANTLDYIDGV----- 654
DB 664 PAARRRSVQLMERRMDKAGQYTDKGLKRCCEBDMKRDIPYSCORARLITQGESCLKA 723
QY 655 -YDNEAVERFMEEN-EGHIV-----DIDFSLGSSPHV-RKHPPEWTW----- 696
DB 724 FMDCCNYITKLAREQRRDHVGLASDVDEDIIPREDISRSHFBSMWTTEBELKEPK 783
QY 697 --LDTNMGYRIYQEEFVTPDSITSWATGVYISEDLGLTTPPELOAFQPFIFLNL 754
DB 784 NGISTKV-----NMIFLKDSITWEILAVLSLSDKGI-CVADPEIETVMQDFIIDRL 835
QY 755 PYSVIRGEFALIEITFNLYKATVEKYIE--KSDKPIILMTSEINATGHOQLTVDS 812
DB 836 PYSVIRNRQVEIRAVLFNY-REOEKLYKVELLHNPAFCSMATAK---RYQTEIETP 890
QY 813 EDGATVLEPIRPTHG--EIPITVALSPTASDAVOMLVYAGIEKXSOSI-LLD-- 867
DB 891 KSSAVVPYIVLKLGLQEVYKAAVFNFIISDVAKILKVPEGRVAKTVAVRTLDE 950
QY 868 -LTDNRLOS---TLKTLSPFPPTVTGSE--VOITAGVGLSPING--IASLIMP 918
DB 951 HLNQGVQREDVNAADLSQVDP---TDBETRILOQTPVQAOMABAVDGEHLKILYV 1007
QY 919 YCGGQNNINFAFNITYILDYLYTKKQOLD--NLKELKALSPRQGYORELLYQREDGFS 975
DB 1008 SGCGQNNIMGMTPTIYLAHYLDQTEQWKEFGLEKQEALELIKKGYTQGLAKPISAVA 1067
QY 976 AFGNVDPSGSTLSAFVLCFLEADPYIDIDONVLRITTYL--KHQKSNGEFMPGRVY 1034
DB 1068 AFNNRPP--STWLTAMWSRFSFLANLAIIDSQVLCGAVKWLILBKOKPDGVFOEDGPYI 1125
QY 1035 HSELQGG---NKSPTVLTAYIVTSLGYR--KTQPNI---DVOESIHFLSEFSRGIS 1084
DB 1126 HOEMLQGRNTEADVSLTAFVIALQEKARDICEQVNSLPISINKAGVYLAST-LNLQ 1184
QY 1085 DNYTLALITVAL---SSVGSPPAKALANLWYRAB--QEGMGQFWVWSSEKLSDSWQPRSL 1140
DB 1185 RPYTVALIGYALALANMKLEBPYLTFLNATAKRNREBEGQQLY----- 1228
QY 1141 DIEVAAYLLSHFLQFQSEGIPIRMWLSKQNSLGGFASSTODTYVALKALSEFAA-LMN 1199
DB 1229 NVEATSYALLALLLKDPGVPPVVRWMLDEREYGGGYSSTATFWFALAQYRADVPD 1288
QY 1200 TERTNIQVTVTGPSSPSPLA-----VVOPMANVINSNGFALCOLANVY 1244
DB 1289 HKDLAMDVSIMHPSASSPTVFRLLWESSGLASBEETKQEGSLTRAKKGQGLTSLVYVY 1348
QY 1245 NVKASGSSRRRSRIONQEAFLDLVAV-----KENKODLNVHDVNVCTSFSGPGRSGMA 1297
DB 1349 HAKVAGKTKCKK-----FDLRVITIKPAPETAKKQDAKSSMIDICTRYIGVDATMS 1401
QY 1298 LMEVNLISGFMPSEALISLSEI---VKKVEYDHCKLN---LYLDSVNET-QPCVNI 1346
DB 1402 ILDISMTMGTFIDTNDLELSSGVDRYISKYEMDKAFSNKNTLIYLEKISHSEBDCLSF 1461

```

```

QY 1347 PAVNFKVSNQDASVSIVDYEPFRQAVRSYNSF-----VKLSGCDLC 1390
DB 1462 KVHGFENGLIOPGSKVYSYTNLESCTRFYHPEKDDGMLSKCHNMC 1511
RESULT 15
ID CO3 MOUSE STANDARD, PRT, 1663 AA.
AC P01027;
DT 21-JUL-1986 (Rel. 01, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Complement C3 precursor (HSE-MSF [Contains: C3a anaphylatoxin]).
GN C3.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxId=10090;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM LONG).
RX MEDLINE=85038854; PubMed=6208565;
RA Fey G.H., Lundwall A., Wetzel R.A., Tack B.F., de Bruijn M.H.L.,
Domey H.;
RT "Nucleotide sequence of complementary DNA and derived amino acid
sequence of murine complement protein C3."
RL Philos. Trans. R. Soc. Lond., B, Biol. Sci. 306:333-344(1984).
RN [2]
RP SEQUENCE OF 671-1663 FROM N.A. (ISOFORM LONG).
RX MEDLINE=85054819; PubMed=6094532;
RA Wetzel R.A., Lundwall A., Davidson F., Gibson T., Tack B.F., Fey G.H.;
RT "Structure of murine complement component C3. II. Nucleotide sequence
of cloned complementary DNA coding for the alpha chain."
RL J. Biol. Chem. 259:13857-13862(1984).
RN [3]
RP SEQUENCE OF 671-748 FROM N.A.
RX MEDLINE=83117730; PubMed=6961437;
RA Domey H., Wiebauer K., Kazmaler M., Mueller V., Odink K., Fey G.H.;
RT "Characterization of the mRNA and cloned cDNA specifying the third
component of mouse complement."
RL Proc. Natl. Acad. Sci. U.S.A. 79:7619-7623(1982).
RN [4]
RP SEQUENCE OF 658-761 FROM N.A.
RX MEDLINE=84201365; PubMed=6609661;
RA Fey G.H., Wiebauer K., Domey H.;
RT "Amino acid sequences of mouse complement C3 derived from nucleotide
sequences of cloned cDNA."
RL Ann. N.Y. Acad. Sci. 421:307-312(1983).
RN [5]
RP SEQUENCE OF 1-34 FROM N.A.
RX MEDLINE=83117622; PubMed=6985486;
RA Wiebauer K., Domey H., Diggelmann H., Fey G.;
RT "Isolation and analysis of genomic DNA clones encoding the third
component of mouse complement."
RL Proc. Natl. Acad. Sci. U.S.A. 79:7077-7081(1982).
RN [6]
RP SEQUENCE OF 25-41 AND 749-760.
RX MEDLINE=93373334; PubMed=8364938;
RA Hamada U.-I., Cavanaugh P.G., Miki K., Nicolson G.L.;
RT "A paracrine migration-stimulating factor for metastatic tumor cells
secreted by mouse hepatic sinusoidal endothelial cells:
identification as complement component C3b."
RL Cancer Res. 53:4418-4423(1993).
RN [7]
RP ALTERNATIVE INITIATION.
RX MEDLINE=95053742; PubMed=7964485;
RA Caban-Kramer Y., Martenson I.L., Melchers F.;
RT "The structure of an alternate form of complement C3 that displays
costimulatory growth factor activity for B lymphocytes."
RL J. Exp. Med. 180:2079-2088(1994).
CC -I- FUNCTION: C3 PLAYS A CENTRAL ROLE IN THE ACTIVATION OF THE
CC COMPLEMENT SYSTEM. ITS PROCESSING BY C3 CONVERTASE IS THE CENTRAL
CC REACTION IN BOTH CLASSICAL AND ALTERNATIVE COMPLEMENT PATHWAYS.
CC AFTER ACTIVATION C3B CAN BIND COVALENTLY, VIA ITS REACTIVE

```

CC THIOLESTER, TO CELL SURFACE CARBOHYDRATES OR IMMUNE AGGREGATES.  
 CC -1- FUNCTION: DERIVED FROM PROTEOLYTIC DEGRADATION OF COMPLEMENT C3,  
 CC C3A ANAPHYLATOXIN IS A MEDIATOR OF LOCAL INFLAMMATORY PROCESS. IT  
 CC INDUCES THE CONTRACTION OF SMOOTH MUSCLE, INCREASES VASCULAR  
 CC PERMEABILITY AND CAUSES HISTAMINE RELEASE FROM MAST CELLS AND  
 CC BASOPHILIC LEUKOCYTES. THE SHORT ISOFORM HAS B-CELL STIMULATORY  
 CC ACTIVITY.  
 CC -1- SUBUNIT: C3 precursor is first processed by the removal of 4 Arg  
 CC residues, forming two chains, beta and alpha, linked by a  
 CC disulfide bond. C3 convertase activates C3 by cleaving the alpha  
 CC chain, releasing C3a anaphylatoxin and generating C3b (beta chain  
 CC + alpha' chain).  
 CC -1- ALTERNATIVE PRODUCTS:  
 CC Event-Alternative initiation;  
 CC Comment-2 isoforms, long (shown here) and short, are produced by  
 CC alternative initiation;  
 CC -1- MISCELLANEOUS: C3b IS RAPIDLY SPLIT IN TWO POSITIONS BY FACTOR I  
 CC AND A COFACTOR TO FORM IC3b (INACTIVATED C3b) AND C3f WHICH IS  
 CC RELEASED.  
 CC -1- MISCELLANEOUS: IC3b IS THE SLOWLY CLEAVED (POSSIBLY BY FACTOR I)  
 CC TO FORM C3c AND C3dg. OTHER PROTEASES PRODUCE OTHER FRAGMENTS SUCH  
 CC AS C3d OR C3g.  
 CC -1- SIMILARITY: TO C4, C5 AND ALPHA-2-MACROGLOBULIN.  
 CC -1- SIMILARITY: Contains 1 anaphylatoxin-like domain.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.ebi.ac.uk/announcements/>  
 CC or send an email to [license@ebi.ac.uk](mailto:license@ebi.ac.uk)).  
 CC -----  
 DR EMBL; K02782; AAC42013.1; -  
 DR EMBL; J00369; AAA37336.1; -  
 DR EMBL; J00367; AAA37336.1; JOINED.  
 DR EMBL; M33032; AAA37378.1; -  
 DR EMBL; Z37998; CAA86099.2; -  
 DR PIR; A92459; C3MS.  
 DR HSP; P01024; 1C3D.  
 DR MGD; MGI:88227; C3.  
 DR InterPro; IPR002890; A2M\_N.  
 DR InterPro; IPR000020; Anaphylatoxin.  
 DR InterPro; IPR001599; Macroglobln2.  
 DR InterPro; IPR001134; Netrin\_C.  
 DR Pfam; PF00207; A2M; 1.  
 DR Pfam; PF01835; A2M\_N; 1.  
 DR Pfam; PF01821; ANATO; 1.  
 DR Pfam; PF01759; NTR; 1.  
 DR ProDom; PD003264; Anaphylatoxin; 1.  
 DR SMART; SM00643; C345C; 1.  
 DR PROSITE; PS00477; ALPHA\_2\_MACROGLOBULIN; 1.  
 DR PROSITE; PS01177; ANAPHYLATOXIN\_1; 1.  
 DR PROSITE; PS01178; ANAPHYLATOXIN\_2; 1.  
 KM Complement pathway: Complement alternate pathway: Plasma;  
 KM Inflammatory response; Glycoprotein; Signal; Alternative initiation;  
 KM Thioester bond.  
 FT SIGNAL 1 24  
 FT CHAIN 25 1663  
 FT CHAIN 25 666  
 FT CHAIN 671 1663  
 FT CHAIN 1129 1663  
 FT INIT MET 1129 1129  
 FT PEPTIDE 671 748  
 FT CHAIN 749 1663  
 FT PEPTIDE 749 954  
 FT PEPTIDE 955 1303  
 FT PEPTIDE 955 1001  
 FT PEPTIDE 1002 1303  
 FT PEPTIDE 1304 1320  
 FT SITE 748 749  
 FT SITE 1303 1304

FT SITE 1320 1321  
 FT DOMAIN 693 728  
 FT DISULFID 559 816  
 FT DISULFID 626 661  
 FT DISULFID 693 720  
 FT DISULFID 694 727  
 FT DISULFID 707 728  
 FT DISULFID 873 1513  
 FT DISULFID 1101 1158  
 FT DISULFID 1358 1489  
 FT DISULFID 1389 1458  
 FT DISULFID 1506 1511  
 FT DISULFID 1518 1590  
 FT DISULFID 1537 1661  
 FT DISULFID 1637 1646  
 FT CARBOHYD 939 939  
 FT CARBOHYD 1617 1617  
 FT CROSSLINK 1010 1013  
 SQ SEQUENCE 1663 AA; 186482 MW; DE5546CCT69BEA19 CRC64;  
 Query Match 13.0%; Score 956; DB 1; Length 1663;  
 Best Local Similarity 24.0%; Pred. No. 2.3e-44;  
 Matches 384; Conservative 286; Mismatches 626; Indels 302; Gaps 66;  
 QY 3 GPPLTAHLICVCTAALAAVAPRPLVAPGIRPGANTIGVLEL--HCPQVTVK 59  
 DB 6 GSQULVLLILLASPLALGI---PMSIIPNVLRLESEITVLENDAGDIPVTVQ 62  
 QY 60 AELLKTASNLTVSLVLE-AEGVEPKSEFVTLTP-SPLNSADIEIYELRVGTODEILFS 117  
 DB 63 DFLKRLQVLTSEKVTLTGASGHLRSVSIK--IPASKEFNDSKGGHKYTVVAVNGETIVE 119  
 QY 118 NSTRLSEFETKRIISVFIOTDKALKPKQOEYKFRIVLTLESDEPKPKTSINTLIDPKSLIQ 177  
 DB 120 KAVWVSFQSGYL--FIQTDKTYITPGSTVLYRIFVDNNLLPVGKTVVILIEPDGIPVK 177  
 QY 178 Q-WLQOSDGLVISTKFPOLSHPILDGWSIQV---QVNDQTVYQSPQVSEYVLPKFEVTL 233  
 DB 178 RDLISSNNGGILPLSNIPELVNMGMOKIRAYEHAKPIFAFEVKEVYLPSEFVRV 237  
 QY 234 Q--TPLYCSNMSKRLNGITAKTYTGKPKVGDVTLFTPLSPFGKKK---NITKTFKI 286  
 DB 238 EPIETFFYIIDDPNGLEVISIAKFLYGNVDGTAFFVFGVD--GDKKISLASHLTRVIE 295  
 QY 287 NGSANFENDEBKKNVDSN-----GLSEYLDLSGEPVILITVYESVTGISRNVSTN 341  
 DB 296 DVGADAVLTTRKVLMEGRPSNADALVGKSLYVS-----VTVILHSGSDMVEARSG 346  
 QY 342 VFFKQHDYIEFPDYTVLKPISNFTATVKTADGNQLTLEBRNNVTVTQRYNTEY 401  
 DB 347 IPIVTSYQIHFHTKTPFFKPMFPDLMTVPNDGSPA-----SKLVVTD----- 393  
 QY 402 WSGSNS-GNGKMEAVQKINTVPOG-----TKIEPPILEDSE---LQLKAYFLGSK 451  
 DB 394 --GSNAKALQDDQVAKLSINTPNSRQPLTIVRTKQDLPESRQATKEAHAPYSTMAN 451  
 QY 452 SSMAVHSLFSPKTYITQLTRDENITVSGSPFELVYSGNRLKELSTMVVSROQLAVVGK 511  
 DB 452 SNNTLH-----SVSRMELRPGD-NLVNPHLRTDPGHEAKIRYTYLVNKKGLVAGR 505  
 QY 512 -----NSTMPSL--TPENSWTPKACVIVY-----IEDGELISDVLKIPVQ-----L 553  
 DB 506 QVREPGQDLVLSLPTPE--FIPSPFLVAYITLIGASGQREVAVDSVWDVADSCIGTL 563  
 QY 554 VPKNKIKLYSKYKABSEKVSRLISVTPDSIVGIVADVKNLNMASNDITMENVHE 613  
 DB 564 VVKGDPR---DNHLARGGQDTLRIRBGNQ--GARVGLAVAVKGVFLNKKKLTQSKIMDV 618  
 QY 614 LELYNTG-----YIIGMFMNSFAVYQ-----EC----- 636  
 DB 619 VEKADIGCTPGSGKNVAGVMDAGLAKTISQGLQTEQRADLECTKPAARRRRSVOLMERR 678



```

QY 637 ----GLMVLTDANLTKDYIDVYD-----NAYAEFMEENE-----GHYVDI-- 675
Db 679 MUKAQY--TDKGRKCCEDGMRDIPMRYSQCRRAFLITQGERNCIKAFIDCCNHITKRE 736
QY 676 ----HDFSLGSS-----PH-----VRGHPEETWIM-----LDTNMGRIYQ 707
Db 737 QHRDHVGLARSELLEEDIIPEEDIIKSHFPOSMLMTIEELKEPKNGISTKV----- 790
QY 708 BEEVTVPDSITSMVATGFIISBDLGLTTPPELOAFQFPFIPLNLPYSVIRGEFALE 767
Db 791 -MNIFLKOSITTWELAVSLDXKGI-CVADPYEIRVMQDFIDLRLPYSVVRNEGEVIR 848
QY 768 ITIFNYLKDATEVKYIIE--KSDKPEILMTSSEINATGHQOTLLVPSDEGATVLPPIRPT 825
Db 849 AYLFFNY-REOEELKYRVELLHNHPAFCSMATAK---NRYFQTIKIPPKSSVAVPYIVPL 903
QY 826 HLG--EIPITVALSPTASDAVTOMLVAKEGIEKYSOSIILDLTD-----NRLQ 874
Db 904 KIGQOEVEVKAAPFNHFISDGVKTKLVPEGR--INTVAIHITLDPEKLGQGVQKVD 961
QY 875 STLKTLFSFPNTVTGSEH---VOITAGDVLGPSING--LASLRMPYCGCEQNMINF 929
Db 962 VPAADLSQVDPD---TDSETRIILOGSPVVOAEDAVDERLKHILVTGACGCEQNMIGM 1018
QY 930 ADNIYILDYLYTKKKQULTD--NLKEKALSFMROGYQRELLYQREDSFSAFGNYPDSGST 986
Db 1019 TPTVIAVHYLDQTEQWEKEFGIEKQFALBELIKKGYTQALAFKOPSSAVAAFNRRP--ST 1076
QY 987 WISAFVLRCELEADPYIDIDQVULHRTYTWL--KGHOKSNGEFMDPGRVHSELQGG--- 1041
Db 1077 WLTAAVVKVFSLAANLIALDSHVLCGAVMWLILEKQKPDGVQEDGPVTHOEMIGGFRNA 1136
QY 1042 NKSPLYLFAVITSLLGYR---KYOPNI---DVOESIHFLSEFSRGISDNTYLIATLYA 1095
Db 1137 KEADVSLTRFVIALQEARDICGQVNSLPGSINKAGEYTEASY-MNLQRPYVALAGYA 1195
QY 1096 L---SSVGSPPAKKALNMLTWRAEOGGMQFVWSSESKLSDSWQPRSLDIEVAAYALLSH 1152
Db 1196 LALMNKLEBPYLGKFLNTAKDRNR-----WEEPDQOL-----YVVEATSYALLAL 1240
QY 1153 FLQFQTSBGITPMRWLSRQNSLGGFASQODTTVALKALSEPA-LMNTERTNIQYTTVG 1211
Db 1241 LLLKQFDSVPVPRWMLNEQRYGGGYSTOATFMVFOALAQYOTDVPDHKDLNMDVSFHL 1300
QY 1212 PSSPSPPLA-----VVOQPMVNI SANGFGAICQIANVYVVKASGSSRRRR 1256
Db 1301 PRRSSATTTFRLLMENGNIIRSEBTKONEBPSLAKGKRGTTLSVAVVYHAKL-----K 1353
QY 1257 SIONQAPDLDAV-----KENKDLNHDVLMNCTSFSGPGRSGMALMEVNLISGFNV 1309
Db 1354 SKVTCCKKFLRVISIRAPETAKKBEAKTWFLEICTKYLGDVDAITMSILDISMTGFAF 1413
QY 1310 PSEAI SL-----SEYVKVVEYDHGKLN-----LYLDSVNET-QFCVNI PAVRNEKYSNTQ 1358
Db 1414 DTKDLELLASGVDRYISKYEMKKA FSNKXLTLLIYLEKISHTEBEDCLTFKHQYFNVGLIQ 1473
QY 1359 DASVSIVDYEPBRQAVRSNSE-----VKLSSCDLC 1390
Db 1474 PGSVKVSYYSTYNEBCTRFYHPEKODGMLSKLCHSEMC 1511

```

Search completed: January 15, 2004, 18:12:59  
 Job time : 31 secs

**THIS PAGE BLANK (b)(7)(D)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compen Ltd.

OM protein - protein search, using sw model

Run on: January 15, 2004, 18:09:19 ; Search time 53 Seconds  
(without alignments)  
6952.814 Million cell updates/sec

Title: US-10-020-095-4

Perfect score: 7348  
Sequence: 1 MGGPPLTAAHLVCVCTAAL.....HSSVIFPCFKLYEMELWL 1428

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

SPTREMBL\_23:\*

- 1: sp\_archaea:\*
- 2: sp\_bacteria:\*
- 3: sp\_fungi:\*
- 4: sp\_human:\*
- 5: sp\_invertebrate:\*
- 6: sp\_mammal:\*
- 7: sp\_mhc:\*
- 8: sp\_organelle:\*
- 9: sp\_phage:\*
- 10: sp\_plant:\*
- 11: sp\_rodent:\*
- 12: sp\_virus:\*
- 13: sp\_vertebrate:\*
- 14: sp\_unclassified:\*
- 15: sp\_virus:\*
- 16: sp\_bacteriap:\*
- 17: sp\_archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	7326.5	99.7	1445	Q8TDJ3	Q8tdj3 homo sapien
2	5360	72.9	1442	Q8t422	Q8t422 mus musculu
3	4305.5	58.6	4	Q8N3A7	Q8n3a7 homo sapien
4	3360	45.7	654	Q8N915	Q8n915 homo sapien
5	1850	25.2	1519	O46015	O46015 caenorhabdi
6	1779.5	24.2	1809	O8T398	O8t398 ciona intes
7	1703	23.2	1388	O8IPHS	O8iphs drosophila
8	1699.5	23.1	1397	O9VLZ0	O9vlz0 drosophila
9	1693	23.0	1408	O8IPH4	O8iph4 drosophila
10	1692	23.0	1420	O9NFW7	O9nfw7 drosophila
11	1688.5	22.3	1399	O9VLX9	O9vly9 drosophila
12	1638	22.3	1507	O01717	O01717 limulus sp.
13	1542.5	21.0	1884	O9ULD7	O9uld7 homo sapien
14	1542.5	21.0	1885	O81ZJ3	O81zj3 homo sapien
15	1491	20.3	1476	O60486	O60486 cavia porce
16	1463.5	19.9	1496	O9NFW5	O9ntf5 drosophila

17	1461.5	19.9	1496	5	O9VIT9	O9vit9 drosophila
18	1460	19.9	1354	5	O9NFW8	O9ntf8 drosophila
19	1455.5	19.8	1494	5	O8T7T6	O8t7t6 onthodora
20	1428.5	19.4	1732	5	O969A4	O969a4 branchiost
21	1425.5	19.4	1469	5	O9NFW6	O9ntf6 drosophila
22	1425.5	19.4	1469	5	O9NFW6	O9ntf6 drosophila
23	1424	19.4	1503	13	O91076	O91076 lampetra ja
24	1417	19.3	1487	11	O03626	O03626 rattus norv
25	1384	18.8	1461	13	O12978	O12978 xenopus lae
26	1384	18.8	1500	11	O63041	O63041 rattus norv
27	1384	18.8	1500	11	O63332	O63332 rattus norv
28	1374.5	18.7	1464	11	O60488	O60488 cavia porce
29	1354	18.4	1269	5	O9NKT5	O9nkt5 drosophila
30	1328.5	18.1	1760	5	O9VLT3	O9vlt3 drosophila
31	1321	18.0	1442	13	O9PVU5	O9pvu5 cyprinus ca
32	1310.5	17.8	1728	5	O817P1	O817p1 swifitia exs
33	1284.5	17.5	1406	13	O9PVU4	O9pvu4 cyprinus ca
34	1233.5	16.8	1340	5	O9GYW4	O9gyw4 anopheles g
35	1105.5	15.0	1699	5	O44344	O44344 strongyloce
36	1099.5	15.0	798	5	O8MT94	O8mt94 drosophila
37	1079.5	14.7	1683	13	O91741	O91741 xenopus lae
38	1073	14.6	1700	13	O91933	O91933 cyprinus ca
39	1054.5	14.4	1489	13	O8QSD4	O8qsd4 xana catesb
40	1054	14.3	1716	13	O91932	O91932 cyprinus ca
41	990	13.5	1652	13	O90633	O90633 gallus gall
42	985	13.4	1738	11	O70346	O70346 mus musculu
43	983	13.4	1614	13	O98977	O98977 oncothynchu
44	965	13.1	1662	13	O98TS6	O98ts6 anathiclas
45	963	13.1	785	11	O62591	O62591 rattus norv

## ALIGNMENTS

RESULT 1	ID	PRELIMINARY	PRT	1445 AA.
Q8TDJ3	Q8TDJ3			
AC	Q8TDJ3			
DT	01-JUN-2002 (TREMBLrel. 21, Created)			
DT	01-JUN-2002 (TREMBLrel. 21, Last sequence update)			
DT	01-OCT-2002 (TREMBLrel. 22, Last annotation update)			
DE	CD109.			
GN	CD109.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=21849742; PubMed=11861284.			
RA	Lin M., Sutherland D.R., Horsfall W., Totty N., Yeo E., Nayar R.,			
RA	Wu X.F., Schun A.C.			
RT	"Cell surface antigen CD109 is a novel member of the alpha(2)			
RT	macroglobulin/C3, C4, C5 family of thioester-containing proteins."			
RL	Blood 99:1683-1691(2002).			
DR	EMBL: AF410459; AAL84159.1; -			
DR	InterPro; IPR002890; A2M_N.			
DR	InterPro; IPR01599; Macrogloblna2.			
DR	Pfam; PF00207; A2M; 1.			
DR	Pfam; PF01835; A2M_N; 1.			
DR	PROSITE; PS00477; ALPHA_2_MACROGLOBULIN; 1.			
SO	SEQUENCE 1445 AA; 161719 MW; 6E8ED2D936AF310D CRC64;			
Query Match	99.7%; Score 7326.5; DB 4; Length 1445;			
Best Local Similarity	98.8%; Pred. No. 0;			
Matches 1427; Conservative	1; Mismatches 0; Indels 17; Gaps 1;			
QY	1 MGGPPLTAAHLVCVCTAALVAPGPRPLVAPGIRGAVNTTGVLEHCPSQVYTKA 60			
DB	1 MGGPPLTAAHLVCVCTAALVAPGPRPLVAPGIRGAVNTTGVLEHCPSQVYTKA 60			
QY	61 ELKTSNLTVSVLEAGVFEKSGFKTLTLPSPINSADETIYELRVTCRTDEILFNSST 120			

```

Db 61 ELKTSANLTVSVLEAGVEFKSFKTLTLPSPLENSADEIYELRVGTODEILFNSNT 120
QY 121 RLSEFKRIISVFIQTDKALYKPKQEVKFRIVTLTSPDEPKYKTLNLIKDPKSNLIQOML 180
Db 121 RLSEFKRIISVFIQTDKALYKPKQEVKFRIVTLTSPDEPKYKTLNLIKDPKSNLIQOML 180
QY 181 SQOSDLGVISKTFOQLSSHPILGDMSIQOVVNDQTYYSQVSEYVLPKFEVTLQTPLYCS 240
Db 181 SQOSDLGVISKTFOQLSSHPILGDMSIQOVVNDQTYYSQVSEYVLPKFEVTLQTPLYCS 240
QY 241 MNSKHLNGITTAKTGYKPKVGDVTLTFLPLSPFKKKNTTKTKINGSANFSNDEEMK 300
Db 241 MNSKHLNGITTAKTGYKPKVGDVTLTFLPLSPFKKKNTTKTKINGSANFSNDEEMK 300
QY 301 NMDSSNGLSSEYLDLSSPGVEILTTVTESTGSRNVSTNVFKQHDYIIIEFDYTVL 360
Db 301 NMDSSNGLSSEYLDLSSPGVEILTTVTESTGSRNVSTNVFKQHDYIIIEFDYTVL 360
QY 361 KPSLNFATVYKTRADGNQTLTERRNNVITVTQRYTEYWSGNSGNQMEAVQKINY 420
Db 361 KPSLNFATVYKTRADGNQTLTERRNNVITVTQRYTEYWSGNSGNQMEAVQKINY 420
QY 421 TVPQSGTFKIEFPILEDSSELQKAYFLGSKSMAVHSLPKSPKTYIOLKTRDENIKYG 480
Db 421 TVPQSGTFKIEFPILEDSSELQKAYFLGSKSMAVHSLPKSPKTYIOLKTRDENIKYG 480
QY 481 SPFLVYVSGNKLKELSTYVVRGOLVAVGKONSTWELPENSMTPKACIYIYIEDDG 540
Db 481 SPFLVYVSGNKLKELSTYVVRGOLVAVGKONSTWELPENSMTPKACIYIYIEDDG 540
QY 541 EIIISDLKIPVQVLFKXKIKLYMSKVAEPSEKYSLSISVTPDPSIVGIVAVDKSVLNM 600
Db 541 EIIISDLKIPVQVLFKXKIKLYMSKVAEPSEKYSLSISVTPDPSIVGIVAVDKSVLNM 600
QY 601 ASNDITMENVHLELVTYGYLLGMPNNSPVPFQCGMLTDMNLTKDYIDGYDIAEX 660
Db 601 ASNDITMENVHLELVTYGYLLGMPNNSPVPFQCGMLTDMNLTKDYIDGYDIAEX 660
QY 661 AERFMEENEGHIVIHDFSLGSSPHVKGHPETIMWIDMNGYIYOEFVTVDSITSM 720
Db 661 AERFMEENEGHIVIHDFSLGSSPHVKGHPETIMWIDMNGYIYOEFVTVDSITSM 720
QY 721 VATGFVISEDGLGLTTPVELOAFOPFIFLNPYSVIRGEFALETIFNYLKADATEV 780
Db 721 VATGFVISEDGLGLTTPVELOAFOPFIFLNPYSVIRGEFALETIFNYLKADATEV 780
QY 781 KVIIEKSDKPDIIIMTSEINATGHQOTLVSEEDGATVLPRIPTHGEIPIVTALSP 840
Db 781 KVIIEKSDKPDIIIMTSEINATGHQOTLVSEEDGATVLPRIPTHGEIPIVTALSP 840
QY 841 ASDAVTOMILVKAIGIEKSYQSILDLTDNRLOSTKTSFSPPTVTGSEKVOITAI 900
Db 841 ASDAVTOMILVKAIGIEKSYQSILDLTDNRLOSTKTSFSPPTVTGSEKVOITAI 900
QY 901 GDVLGPIINGLASLIRMPYGGGEONMINFAPNIYILDYLTKKQJLTNLEKALSPRQ 960
Db 901 GDVLGPIINGLASLIRMPYGGGEONMINFAPNIYILDYLTKKQJLTNLEKALSPRQ 960
QY 961 YQRELLVQREDESSAGANDPSSGTWLSAFVLCFLEADPYIDIDONVLRHTYTWLKG 1020
Db 961 YQRELLVQREDESSAGANDPSSGTWLSAFVLCFLEADPYIDIDONVLRHTYTWLKG 1020
QY 1021 OKSNGEPMDEGRVHSELOGNKSPTVLTAYITVSLGYRYQONIDVOESIHLESEFS 1080
Db 1021 OKSNGEPMDEGRVHSELOGNKSPTVLTAYITVSLGYRYQONIDVOESIHLESEFS 1080
QY 1081 RGISDNVTALITVLSVSPKAKALNMLTPRABEGWQFVWSESKLSDWQPRSL 1140
Db 1081 RGISDNVTALITVLSVSPKAKALNMLTPRABEGWQFVWSESKLSDWQPRSL 1140
QY 1141 DIEVAAYALISHFOFOTSEGIPIMRWLSRORNSLGSPASTODTTVALKALSEFALMNT 1200
Db 1141 DIEVAAYALISHFOFOTSEGIPIMRWLSRORNSLGSPASTODTTVALKALSEFALMNT 1200

```

```

QY 1201 ERTNIQVTVGPSSPSP-----LAVOPMAVNISANGFPAICOLNV 1243
Db 1201 ERTNIQVTVGPSSPSPKFLDTHNRLLQTAELAVOPMAVNISANGFPAICOLNV 1260
QY 1244 YNVKASGSSRRRSIONOEAFLDVAVKENKDDLNVLDVCTSBEGPGRSGMALMEVNL 1303
Db 1261 YNVKASGSSRRRSIONOEAFLDVAVKENKDDLNVLDVCTSBEGPGRSGMALMEVNL 1320
QY 1304 LSGFMVPSAISISETVKKVEYDHGKLNLYLDSVNETOPCVNIPAVRNFKVSTODASYS 1363
Db 1321 LSGFMVPSAISISETVKKVEYDHGKLNLYLDSVNETOPCVNIPAVRNFKVSTODASYS 1380
QY 1364 IVDYEPRRQAVASYNSEVYLSGCDLCSVOGCRPCEDGASGSHHSYIIFCFKLXYF 1423
Db 1381 IVDYEPRRQAVASYNSEVYLSGCDLCSVOGCRPCEDGASGSHHSYIIFCFKLXYF 1440
QY 1424 MELML 1428
Db 1441 MELML 1445

RESULT 2
QY 08422 PRELIMINARY; PRT; 1442 AA.
ID 08422;
AC 08422;
DT 01-JUN-2002 (Tremblrel. 21, Created)
DT 01-JUN-2002 (Tremblrel. 21, last sequence update)
DT 01-OCT-2002 (Tremblrel. 22, last annotation update)
DE GPI-anchored alpha-2 macroglobulin-related protein.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Hashimoto M., Ichihara M., Takahashi M.;
RT "Cloning and characterization of GPI-anchored alpha-2 macroglobulin-
RT related protein.";
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL: AY083458; AAM00021.1; -.
DR InterPro: IPR002890; A2M_N.
DR InterPro: IPR001599; Macroglloblna2.
DR Pfam: PF00207; A2M; 1.
DR Pfam: PF01835; A2M_N; 1.
DR PROSITE: PS00477; ALPHA 2 MACROGLOBULIN; 1.
SQ SEQUENCE 1442 AA; 161658 MW; E2B9671BBEASBA7 CRC64;

Query Match 72.9%; Score 5360; DB 11; Length 1442;
Best Local Similarity 71.6%; Pred. No. 4,1e-318;
Matches 1038; Conservative 176; Mismatches 202; Indels 34; Gaps 8;

QY 1 MGPSPILTAHLICVCTALAAVAPGRPLVTAGIIRRGNAVITIGVLEHCHPSQVTVRA 60
Db 1 MGRSRILISAHLICCAVALA-APGSRPLVTAGIIRPGNAVITIGVLDLNSPPQVLVA 59
QY 61 ELKTSANLTVSVLEAGVEFKSFKTLTLPSPLENSADEIYELRVGTODEILFNSNT 120
Db 60 QVTKINSKRSRSLIEBGFPHRHFVTLVPLPLSSAKITELHNGOSEMEIVSNRT 119
QY 121 RLSFETKRISVFIQTDKALYKPKQEVKFRIVTLTSPDEPKYKTLNLIKDPKSNLIQOML 180
Db 120 RLTFESKSIISVLIQTDKAFKPKQEVKFRIVTLTSPDEPKYKTSVDIFIDPKSNVLIQOM 179
QY 181 SQOSDLGVISKTFOQLSSHPILGDMSIQOVVNDQTYYSQVSEYVLPKFEVTLQTPLYCS 240
Db 180 SQGDDGVVSKTFOQLSSNPIFGDMSIQOVVNDQTYYSQVSEYVLPKFEVTLQTPLYCS 239
QY 241 MNSKHLNGITTAKTGYKPKVGDVTLTFLPLSPFKKKNTTKTKINGSANFSNDEEMK 300
Db 240 LNSKOLNGSVIATKTYTGKPEVKSLSLTPSPFKKKNTKTSFELNGRANFSFDVYEMK 299
QY 301 NVM-----DSSNGLSSEYLDLSSPGVEILTTVTESTGSRNVSTNVFKQHDYIIIEF 354

```

Db 300 KWNMLKPLTDBSSEYENDPSPGPBAEIAITVTESLTGISRMASTNVFCKHDYIEIF 359  
 Qy 355 DYTTLVKPSLNFATATVKTADGNOLTLBERNNVITVTO--NYTEWSSGNSGNOM 412  
 Db 360 DYTTLVKPSLNFATATVKTADGNOLTLBERNNVITVTO--NYTEWSSGNSGNOM 414  
 Qy 413 EAVOKINTVPOSGTFKIEFPILDESSLOLKAAYELGSKSMAVHSLFKSPKTYIOLNT 472  
 Db 415 DYTTLVKPSLNFATATVKTADGNOLTLBERNNVITVTO--NYTEWSSGNSGNOM 474  
 Qy 473 RBNENKVSFPFELVSGNKLKELSTMVSRGOLVAVGQNSTMSLTPEBNTWTPACVY 532  
 Db 475 RBNENKVSFPFELVSGNKLKELSTMVSRGOLVAVGQNSTMSLTPEBNTWTPACVY 534  
 Qy 533 VYIEDDEGIIISDVLTPIQVLFKPKIKLYMSKVKAPESEKSLRISVYOPISYIVAV 592  
 Db 535 AYIIMADEGIIINDLKIPIQVLFKPKIKLYMSKVKAPESEKSLRISVYOPISYIVAV 594  
 Qy 593 DKSVMIMNANDITMENVHELELYNTGYILGFMMNSFAVPOEGGLMVLTDANLTEDYD 652  
 Db 595 DKSVMIMNANDITMENVHELELYNTGYILGFMMNSFAVPOEGGLMVLTDANLTEDYD 654  
 Qy 653 GYIDNAEYAKERMENEGHIVDIDFSLGSSPHVRKHPEPTWIMLDTNNGYRIYOEFEV 712  
 Db 655 EYIDTEEYSEKPEAEENAMLVDPEDASVANNVHRKNPEPTWIMLDTNNGYRIYOEFEV 714  
 Qy 713 VDDSTISWATGVFVSEDGLGLTTTPVLOAPOPFFIPLNIPYSYIRBEEFALETTYN 772  
 Db 715 VDDSTISWATGVFVSEDGLGLTTTPVLOAPOPFFIPLNIPYSYIRBEEFALETTYN 774  
 Qy 773 YLKDATVVKVIEKSDKFDILMTSEINATGHOOTLVSEBDGATVLPFRPHLGEIP 832  
 Db 775 YLKDATVVKVIEKSDKFDILMTSEINATGHOOTLVSEBDGATVLPFRPHLGEIP 834  
 Qy 833 TTTALSPASDAVTOQMLVKAEGIEKSYQSILDLTDNRLOSTLKTLSFSPPTVTGS 892  
 Db 835 TTTALSPASDAVTOQMLVKAEGIEKSYQSILDLTDNRLOSTLKTLSFSPPTVTGS 894  
 Qy 893 EKVQITAGIDVIGPISNGLASLIRMPYGCGEQNMNINPANIYILDVLTCKKQITDNLK 952  
 Db 895 EKVQITAGIDVIGPISNGLASLIRMPYGCGEQNMNINPANIYILDVLTCKKQITDNLK 954  
 Qy 953 ALSFMRQYORELLVYOREDEFSFAPGNVDPGSGTWLSAFVLRGFLPADYIIDQVLR 1012  
 Db 955 ALSFMRQYORELLVYOREDEFSFAPGNVDPGSGTWLSAFVLRGFLPADYIIDQVLR 1014  
 Qy 1013 TTTMLKQKSGNPFWDPGRVIHSELQGNKSPVTLTAYIVTSLGKRYQOPNIDVQSI 1072  
 Db 1015 TTTMLKQKSGNPFWDPGRVIHSELQGNKSPVTLTAYIVTSLGKRYQOPNIDVQSI 1074  
 Qy 1073 HFLSEFSFGISDNTLTALITVALSVSPKAKAELNMLTWRABOEGMOPVWSSSKL 1132  
 Db 1075 HFLSEFSFGISDNTLTALITVALSVSPKAKAELNMLTWRABOEGMOPVWSSSKL 1134  
 Qy 1133 DSMOPSLDIEVAAYALSHFLQFOTSEGIPIIMRWLSROBNSLGGFASQODTVALKALS 1192  
 Db 1135 DSMOPSLDIEVAAYALSHFLQFOTSEGIPIIMRWLSROBNSLGGFASQODTVALKALS 1193  
 Qy 1193 BEPALMTERINIQTVTGPPSSP-----SPLAVQPMANVISANGFG 1235  
 Db 1194 BEPALMTERINIQTVTGPPSSP-----SPLAVQPMANVISANGFG 1253  
 Qy 1236 AICQALNVVNVKASGSSRRRSIQNOEAPDLVAKENKODLNVDLVNCTSPSGRSG 1295  
 Db 1254 AICQALNVVNVKASGSSRRRSIQNOEAPDLVAKENKODLNVDLVNCTSPSGRSG 1312  
 Qy 1296 MALMEVNLISGFVPEASISLSEFVKVYEDHGXNLVYDSVNEPOFCNIAVANEKFS 1355  
 Db 1313 MALMEVNLISGFVPEASISLSEFVKVYEDHGXNLVYDSVNEPOFCNIAVANEKFS 1372  
 Qy 1356 NTQDASVSIIVDYPERROAVRSYNSFVKLSCDLCSVOGCRPCEDGASGSHHSVIRI 1415

Db 1373 NIRDGSVSMVDYPERROAVRSYNTQVKLSGCCYLSPT-NCKSHTDGATSLRRSSILLV 1431  
 Qy 1416 FCFKLLYFME 1425  
 Db 1432 FCSVLLYFVQ 1441  
 RESULT 3  
 ID 08N3A7 PRELIMINARY; PRT, 854 AA.  
 AC 08N3A7; 01-OCT-2002 (Tremblrel. 22, Created)  
 DT 01-OCT-2002 (Tremblrel. 22, Last sequence update)  
 DE 01-MAR-2003 (Tremblrel. 23, Last annotation update)  
 DT Hypothetical protein (Fragment).  
 GN DKEP762111.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 OX NCBI\_Taxid=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Melanoma;  
 RA Blum H., Baurerachs S., Mewes H.W., Well B., Wiemann S.;  
 RL Submitted (JUL-2002) to the EMBL/Genbank/DBJ databases.  
 DR EMBL; AL834478; CAD39137.1; -;  
 DR InterPro: IPR001599; Macroglolbln2.  
 DR Pfam: PF00207; A2M; 1.  
 DR PROSITE: PS00477; ALPHA\_2\_MACROGLOBULIN; 1.  
 KW Hypothetical protein.  
 FT NON TER 1  
 SQ SEQUENCE 854 AA; 95599 MW; 7B51863A2565C004 CRC64;  
 Query Match 58.6%; Score 4305.5; DB 4; Length 854;  
 Best Local Similarity 97.7%; Pred. No. 4.7e-254;  
 Matches 834; Conservative 1; Mismatches 2; Indels 17; Gaps 1;  
 Qy 592 VDKSVNLNANSDITMENVHELELYNTGYILGFMMNSFAVPOEGGLMVLTDANLTCKYI 651  
 Db 1 VDKSVNLNANSDITMENVHELELYNTGYILGFMMNSFAVPOEGGLMVLTDANLTCKYI 60  
 Qy 652 DGVDNAEYAKERMENEGHIVDIDFSLGSSPHVRKHPEPTWIMLDTNNGYRIYOEFEV 711  
 Db 61 DGVDNAEYAKERMENEGHIVDIDFSLGSSPHVRKHPEPTWIMLDTNNGYRIYOEFEV 120  
 Qy 712 TVPDSITISWATGVFVSEDGLGLTTTPVLOAPOPFFIPLNIPYSYIRBEEFALETTYN 771  
 Db 121 TVPDSITISWATGVFVSEDGLGLTTTPVLOAPOPFFIPLNIPYSYIRBEEFALETTYN 180  
 Qy 772 NYLKDATVVKVIEKSDKFDILMTSEINATGHOOTLVSEBDGATVLPFRPHLGEIP 831  
 Db 181 NYLKDATVVKVIEKSDKFDILMTSEINATGHOOTLVSEBDGATVLPFRPHLGEIP 240  
 Qy 832 ITTVALSPASDAVTOQMLVKAEGIEKSYQSILDLTDNRLOSTLKTLSFSPPTVTGS 891  
 Db 241 ITTVALSPASDAVTOQMLVKAEGIEKSYQSILDLTDNRLOSTLKTLSFSPPTVTGS 300  
 Qy 892 SERVQITAGIDVIGPISNGLASLIRMPYGCGEQNMNINPANIYILDVLTCKKQITDNLK 951  
 Db 301 SERVQITAGIDVIGPISNGLASLIRMPYGCGEQNMNINPANIYILDVLTCKKQITDNLK 360  
 Qy 952 KALSMRQYORELLVYOREDEFSFAPGNVDPGSGTWLSAFVLRGFLPADYIIDQVLR 1011  
 Db 361 KALSMRQYORELLVYOREDEFSFAPGNVDPGSGTWLSAFVLRGFLPADYIIDQVLR 420  
 Qy 1012 RYTYLTKQKSGNPFWDPGRVIHSELQGNKSPVTLTAYIVTSLGKRYQOPNIDVQSI 1071  
 Db 421 RYTYLTKQKSGNPFWDPGRVIHSELQGNKSPVTLTAYIVTSLGKRYQOPNIDVQSI 480  
 Qy 1072 HFLSEFSFGISDNTLTALITVALSVSPKAKAELNMLTWRABOEGMOPVWSSSKL 1131  
 Db 481 HFLSEFSFGISDNTLTALITVALSVSPKAKAELNMLTWRABOEGMOPVWSSSKL 540

```

QY 1132 SDSWQPSRLDIEVAAYALLSHFLQFQTSSEGIPIRMWLSRGRNSLGGFASDTDTVALKAL 1191
DB 541 SDSWQPSRLDIEVAAYALLSHFLQFQTSSEGIPIRMWLSRGRNSLGGFASDTDTVALKAL 600
QY 1192 SEPALMTERNTNQVVTGSSPSPP-----LAVVQPMVNTSANGFG 1234
DB 601 SEPALMTERNTNQVVTGSSPSPPKELIDTNHRLLLQTAELAVVQPMVNTSANGFG 660
QY 1235 FAICQLNVVNVKASGSSRRRRSIOHQEAFDLVAVKENKDDLHVNDLVNCTSPSGPRS 1294
DB 661 FAICQLNVVNVKASGSSRRRRSIOHQEAFDLVAVKENKDDLHVNDLVNCTSPSGPRS 720
QY 1295 GMAIMEVNLISGFVWPSEALISLSETVKKVEYDHGKLNLYDSVNETQFCVNI PAVRNPKV 1354
DB 721 GMAIMEVNLISGFVWPSEALISLSETVKKVEYDHGKLNLYDSVNETQFCVNI PAVRNPKV 780
QY 1355 SNTQDASVSYDYVEPRRQAVRSYNSVEYKSSCULCSVQCCRCEDGASGSHHSSVIF 1414
DB 781 SNTQDASVSYDYVEPRRQAVRSYNSVEYKSSCULCSVQCCRCEDGASGSHHSSVIF 840
QY 1415 IFCEFLTYFMELML 1428
DB 841 IFCEFLTYFMELML 854

RESULT 4
QY 08N915 PRELIMINARY; PRT; 665 AA.
AC 08N915.
DT 01-OCT-2002 (TREMBlrel. 22, Created)
DT 01-OCT-2002 (TREMBlrel. 22, Last sequence update)
DE 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OC NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Oshima A., Takahashi-Fujii A., Tanase T., Imose N., Takeuchi K.,
RA Arita M., Muraashino K., Yuuki H., Hara H., Sugiyama T., Irie R.,
RA Otsuki T., Sato H., Wakamatsu A., Ishii S., Yamamoto J., Isono Y.,
RA Kawai-Hio Y., Salto K., Nishikawa T., Kimura K., Yamashita H.,
RA Matsuo K., Nakamura Y., Sekine M., Kikuchi H., Kanda K., Magatsuma M.,
RA Murakawa K., Kanehori K., Sugiyama A., Kawakami B., Suzuki Y.,
RA Sugano S., Nagahori K., Maehiro Y., Nagai K., Isogai T.;
RT "NEDO human cDNA sequencing project.";
RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK095888; BAC04642.1; -
DR InterPro; IPR002890; A2M_N.
DR Pfam; PF01835; A2M_N; 2.
KM Hypothetical protein.
SQ SEQUENCE 665 AA; 74411 MW; CC73982C8E88647E CRC64;

Query Match 45.7%; Score 3360; DB 4; Length 665;
Best Local Similarity 100.0%; Pred. No. 1.6e-196;
Matches 657; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 241 MNSKHLNGITTAATYTGKPVKGDVTLFLPLSPFGKKKNTTKTEKINGSANFSFNDSEMK 300
DB 241 MNSKHLNGITTAATYTGKPVKGDVTLFLPLSPFGKKKNTTKTEKINGSANFSFNDSEMK 300
QY 301 NVMDSSNGLSSEYLDSPGVEVLTITVTSVGTGSRNVSINVPKQHDYIIIEFDYTVL 360
DB 301 NVMDSSNGLSSEYLDSPGVEVLTITVTSVGTGSRNVSINVPKQHDYIIIEFDYTVL 360
QY 361 KPSLNFATYKVRACDNOQLTEERRNNVITYTORNYTEYMGSGNSGNQKMEAVOKINY 420
DB 361 KPSLNFATYKVRACDNOQLTEERRNNVITYTORNYTEYMGSGNSGNQKMEAVOKINY 420
QY 421 TVPQSTFKIEPFILEDSSSELOKAYFLGSKSSMAVHSLFKSPSKTYIQLKTRDENIKVG 480
DB 421 TVPQSTFKIEPFILEDSSSELOKAYFLGSKSSMAVHSLFKSPSKTYIQLKTRDENIKVG 480
QY 481 SPPELVVSGNKRLEKESYMYVSRGQLVANGKONSIMFSLTPENSWTPKACVIYYIEDDG 540
DB 481 SPPELVVSGNKRLEKESYMYVSRGQLVANGKONSIMFSLTPENSWTPKACVIYYIEDDG 540
QY 541 EITSDVLTKEIPVOLVFNKIKLYMSKVKAPESEKVSRLISTOPDSITVGIYAVDKSVNLN 600
DB 541 EITSDVLTKEIPVOLVFNKIKLYMSKVKAPESEKVSRLISTOPDSITVGIYAVDKSVNLN 600
QY 601 ASNDITMENVHLELVNTGYLGMPNNSFAVFOEGGLMWLTDAINTKDYIDGVYDN 657
DB 601 ASNDITMENVHLELVNTGYLGMPNNSFAVFOEGGLMWLTDAINTKDYIDGVYDN 657

RESULT 5
QY 046015 PRELIMINARY; PRT; 1519 AA.
AC 046015.
DT 01-JUN-1998 (TREMBlrel. 06, Created)
DT 01-JUN-1998 (TREMBlrel. 06, Last sequence update)
DE 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
OS Hypothetical 168.4 kDa protein ZK337.1.
GN ZK337.1.
OC Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditioidea;
OC Rhabditidae; Pelodierinae; Caenorhabditis.
OC NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RA White S.;
RL Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.
CC -1- ALTERNATIVE PRODUCTS: 2 ISOFORMS; ZK337.1B (SHOWN HERE) AND
CC ZK337.1A; ARE PRODUCED BY ALTERNATIVE SPLICING.
DR EMBL; Z82090; CAB05007.1; -
DR EMBL; Z82090; CAB05006.1; -
DR HSP; P01024; 1C3D.
DR WormPep; ZK337.1a; CE16718.
DR WormPep; ZK337.1b; CE16719.
DR InterPro; IPR000847; A2M_N.
DR InterPro; IPR000847; HTH_LySR.
DR InterPro; IPR001599; Macroglublna2.
DR InterPro; IPR001517; Nhrtransf_2.
DR Pfam; PF00207; A2M; 1.
DR Pfam; PF01835; A2M_N; 1.
DR PROSITE; PS00599; AA_TRANSFER_CLASS_2; 1.
DR PROSITE; PS00477; ALPHA_2_MACROGLUBLIN; 1.
DR PROSITE; PS00044; HTH_LYSR_FAMILY; 1.
KM Alternative splicing, Hypothetical protein.
FT VARSPLIC 714 724 MISSING (IN ISOFORM ZK337.1A).
SQ SEQUENCE 1519 AA; 168385 MW; 78139ACB564AB742 CRC64;

Query Match 25.2%; Score 1850; DB 5; Length 1519;
Best Local Similarity 31.6%; Pred. No. 8.6e-104;
Matches 477; Conservative 290; Mismatches 563; Indels 180; Gaps 44;

```



28 VSTTAPVKA-TYMLVAVRPPQPSVCMNLKQATDEDMIVRIEVERTNETIAR 86  
 75 EAEVPEKSGFKTLTLPSPILNSAD--EYELRVTRTOD-EILPSNSTRISFETKRI 131  
 87 VLSNL-KPGIAGTVLSLSEMPASQSLTPROQYKLYIGETINAILIENENELKYDOKALSV 145  
 132 FIOTDKALYKROEVKRIYVTLFSPFKPKYKSLNLIKDPKSNLIQOMLSQSDLGIVSK 191  
 146 FIOTRAIRPAPSLVRYRAIVKSDLKPYGNATIKIFDPSRNLISQITGLTDRGVSG 205  
 192 TQOLSHPIEDGMSIQO---VNDQTYQSFQVSYUYPKFEVTLQTFYSQMSKHLN 247  
 206 ELOLEETLLGDMFLEVESNGVODS---SFTVDTYVLPKFEVNIKTSSFTIIND-DLS 261  
 248 GTTAKYTYGKPVKGDVTLTF-LPLSFV-----GKKKN-----ITKFKIN--GS 289  
 262 VAVDAKTYTGKGVAKAKVSLPHEHRIAMPPTIIDENGVKKEELMERIVKLNROCE 321  
 322 AAVVFSNDELK-----HKLHMGWGGSIIRIVASVTEITELERNATHQISFREV 373  
 350 IIEFDYTVLKPISLNTATKVTADGN--QTLBERNNVITYTQNTYEGSGNS 407  
 374 KLDVEKOGDTFKRGTYNVVALKQMDTPPVKATLPKR---VOVSTFYNP--YNHDS 427  
 408 GNQKMEAVOKINYTVPOSGTFKIBPFILEDSELOKAVFLGSKSMVHSLFK----- 461  
 428 SLQEKETKIVEVDAHGISVLTLOPINCTSARIEAH-YDIGKONFTATPIYSSLYEA 486  
 462 --SPSKTYIQLKTRDEN-IKVGSPELVVSGNKRLELSYMYVSRGOLV-----AVGKON 513  
 487 AVSPKTSFIQLADNAGAVDVGKSLSFSLKATOPSTITYQWMSNSNIYVSOQMTVNSH 546  
 514 STMEFUTPENSWTPEACVIVY-YIEDDEIISDLKIPQVLPKNIKIKLYMSKVAEBS 572  
 547 AIT-SFPATANMAPKSRILIVAIIESSQVLDALDFKEGIFQOVALISIKQAVEPQO 605  
 573 KVSLSISTVQPSIYIVAVDSKVNLMNSNDITMENVHELEIVNTGYGLMF----- 626  
 606 NKFKVY-SDKNSFGLVAVDSDVLLKTGNDITREKVKODLENDYDNNVGGFGGPRPW 664  
 627 -----MNSFAVQECGLWLTDA-----NLT 647  
 665 EALDRKKSIMRPMWGISGSDAQSIFSNAGLVLTDLALYRREBQREMERLANTPGILT 724  
 648 KDYIDGVDAVNAEYERFMEENEGHIVDHFSLGSS--PHVRKHPETWILDTNMGR 704  
 725 VMMMGAPGMAEA--FAAPPMG-----GSSPPPTVRKFPPTHWIMSDIN-STS 771  
 705 IYOEFEVVPDSITSVAVGFISEDLGLTTTPVELQAPQPFITLPPSVIRGEER 764  
 772 GEVEHEIEPDTITSVASTFAINEENGVAPTTSKLVFPFPFIQLNPAVVRGEK 831  
 765 ALEITIFNLKATEVAKVII--EKSDPKPILMTSRI--NATGHOQTLVSEDCAT-- 817  
 832 ALVIVENYMEKEQDVTYTKDKDSGYLLKKDGVVARDEVQGNRIYVAGAGGTSK 891  
 818 -VLPIRPTHLGEIPITVTALSPASDAVOTMLVKAEGIEKSYQSOSILLDTNRLQST 876  
 892 AVYFPFIVPSIGEIPHIAISAGGDAVENMLRVDPQYKVDRLNPIVITLNNSSDRS 951  
 877 LKTISSFPNVTGSEKQVITAGIVLGPSINGLASLIRMYGGCEOMINIPARIYTL 936  
 952 -KNLEIIMVNDVVDGQKARLVITGDMGPVNLNAHKLQMPYGCCEOMLMLVPIILAV 1010  
 937 DYLTKKQOLTMLEKALSPFMQGYORELLYOREDSFAPGVNPDGSGTSLSAFLRCE 996  
 1011 KILBANNRESOLETAKIFIEGQIURELTYKRADNSFSAFGDSKAGSTWLTALFVRSE 1070  
 997 LEADPIYIDQVNLHRTYTLKGHQKSNCEPMDPGVHISELQGG-NKSPVTLTAVYTS 1055  
 1071 HHAQVAFVDPNVISRAVAFILNSQOMESGAFBERGVEHHKDMQGGADGVALTAFVLIS 1130

1056 LIGYKQPNIDVOESIHFLSESRGISDNVYTLALTALYSVSGSPRAKALNMLTWR- 1114  
 1131 IL-----ENGNGEAGVATYLEKHIDVSGNAYTAVAVAYALQAKSKAGAFELKCHK 1185  
 1115 -AEDEGKQFVWSE--SKLSDS---WQPSLDIEVAAYALLSHLOFQTSSEGIPIRM 1167  
 1166 IVEKSGDYKFPASQKVEKLEKESRAYMFPARPVIDETTSYAVLSYLAQNGTSESLIRW 1245  
 1168 LSRQNSLGGFASQDTTVALKALSEFALMTERTNIOYV--TPSSPS-----PLAY 1220  
 1246 LVSQNELGFTSTQDTYMALQALSSYAAVYSDHNSQVTLNGKTHSPDINIRNAIV 1305  
 1221 VQPM-----AVNISANGFPAICQLNVVNVNVAAGSSRRRSRSTIONEAPDLDAVKN 1273  
 1306 LQSYLSSLNDASVINANGTVFPAQLSYST-----YRSLINDAPFCQSEIHEI 1356  
 1274 KDLNHRVDLNVCTSPSGRSGMALMEVNLGFWPSEALSLSETVK-----KVEYDHG 1328  
 1357 RAG-NRLQDLCNVYTRGKSNMALAIDALSGRFAEQVHTLSTIEDLRVEMEKD 1415  
 1329 KLNLYDSVNETQFCVNIPAVRNKRVSNTQDASVYDYEPKQAVASYSYKLSGCD 1388  
 1416 KKNVYFNLGGRPVCLSYSDVTYQVADOKPAMFRLVDYDPEBOLKMTYAKQTRSLQ 1475  
 1389 LCSDVQGRP 1398  
 1476 KCG-EDCW 1483

RESULT 6  
 08T398 PRELIMINARY; PRT; 1809 AA.  
 AC 08T398;  
 DT 01-JUN-2002 (TREMBLrel. 21, Created)  
 DT 01-JUN-2002 (TREMBLrel. 21, Last sequence update)  
 DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)  
 DE Alpha-2-macroglobulin homologue.  
 GN A2M.  
 OS Clona intestinalis.  
 OC Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea; Enterogona;  
 OC Phlebobranchia; Cloniidae; Clona.  
 OX NCBI\_TaxID=7719;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Hepatopancreas;  
 RA Hammond J.A., Nakao M., Yano T., Kemp G.D., Smith V.J.;  
 RT "Complete cDNA of an alpha2 macroglobulin homologue with a catalytic  
 histidine residue from Clona intestinalis."  
 RL Submitted (FEB-2002) to the EMBL/Genbank/DBJ databases.  
 DR EMBL, AJ431688; CAD24311.1; --  
 DR InterPro; IPR002890; A2M\_N.  
 DR InterPro; IPR001064; Crystallin.  
 DR InterPro; IPR001599; Macrogloblna2.  
 DR InterPro; IPR000884; TSPL.  
 DR Pfam; PF00207; A2M; 1.  
 DR Pfam; PF01835; A2M\_N; 1.  
 DR SMART; SM00209; TSPL; 1.  
 DR PROSITE; PS00477; ALPHA\_2\_MACROGLOBULIN; 1.  
 DR PROSITE; PS00225; CRYSTALLIN\_BETAGAMMA; 1.  
 DR PROSITE; PS50092; TSPL; 1.  
 SQ SEQUENCE 1809 AA; 199172 MW; E300DC643946E454 CRC64;

Query Match 24.2%; Score 1779.5; DB 5; Length 1809;  
 Best Local Similarity 30.0%; Pred. No. 2.3e-99;  
 Matches 453; Conservative 293; Mismatches 586; Indels 179; Gaps 41;

13 LCVCTAALAAVAPRGPFVLTARGIIRPGGNVITIGVLEHCSQVTVKALELKTASNLVVS 72  
 16 LIATLSSLATASNV-YNVIYFKHIRPGFNISFTALIDN-ENTVOIHTAFRSMDSFHD 73  
 73 VLEAGVEKSGFKTLTLPSPILN-SADEIYELRVTR--TODEILFSNSTRISFETKR 128

74 STDS---VNSGSSSRISNMNGLPHYSGSHGELNITGTDLVTAGOLFENSTDFQAKS 130  
 QY 129 ISVEIQTDKALYKKEOEKFRIVTLFSPDFKXTSLNLLDKPSNLIQOMLSQOSDLGV 188  
 Db 131 ISILIOTDKALYQGHYTKFRALAKLPDLKLOQSNISITFPDPRGNVMEPEPLAHGV 190  
 QY 189 ISKTFOLSSHPILIDWMSIQOVNDQTYQSFQVSEYVLPEKEVTLQTPLYCSNNSKHLNG 248  
 Db 191 AGGCFSLTKDVAAMKMKVEFMABGFKESLSVEKRYKLPEKVEVKA.PSYIHPQSTGLTI 250  
 QY 249 TITAKTYKGRVKGDVTLTLF-----PL--SFMKC-----KNITKTF-KINGSANFS 293  
 Db 251 KLDKMYTFKGVQGGGLEEVGGYQVPYHGFGRFAPRPTQKKITRRYPNFGYELL 310  
 QY 294 FNDEMKVNDMSNGLSSEYDLSSPGVEILTTVTESVTGSRVSTNVFPEKQDYIEF 353  
 Db 311 ITNEINEEL-GMGASGESI-----ITVGSYTEALTRAPFDQRIADKTNVAVET 362  
 QY 354 PDYTVTLKPSLNTATYKVRADGNQTLTEER-RNNVITVTOBN----- 397  
 Db 363 LVKELTITKPKLKYSAVYIQITEVDKPLPEDRLANNLLNIEYRPRGEPEPTNTVTST 422  
 QY 398 -YTEVYSGNSGNGKMEAVQKINTVPOSGTFKIEPILBDS-SELQKAFGLSKSMA 455  
 Db 423 WYAYRMEETRV-----FVIPSGLVKVITIDAPSDTFTSINFRPY--TNATMS 467  
 QY 456 VHSILFK-----SPSKTYIQLKTRDENIKVSGPELVVSGNKLKELSVVVSQOLVA 508  
 Db 468 QRMALQMTARADBPNSYLOITTEBNSVVGPNMATYIIRTEAVSEFTLIIIRGELIS 527  
 QY 509 VGR-----QNSTMSLTPENSWTPKACVIVYIIEDDGEIISDLKIPQVLVKKR 560  
 Db 528 ERKFTQLSGVENSHLFEFSVEYDMIPGVYLASVYRDEGEIVADYIKLVTALLENQVS 587  
 QY 561 LYNKVKVABEPEKSLRISTVQPSYIGVAVNDKSVNLMAANDITMENVHELELVNTG 620  
 Db 588 ITSSSTNIDAGEDVSIRVQTSSSGAYVGARIDSVLLKSGNVOSERITVDLTKYSVT 647  
 QY 621 YLLCMF-----NNSFAVFOECGLMVLTDANLTGDYIDGYD-----NAE 659  
 Db 648 QELNHRMRRMWWYTPPSGASDASDVFRKAGILVFTDLVYKPEASISIPPRPIAFSLNG 707  
 QY 660 YAEFMEENEGHVIDH--DFSLGSSP-HVRKHPETWIMLDTNMGRYIQEPEVTVDS 716  
 Db 708 FAER-----NIATAVNDTSTPATPTKRTLTLPETLMDEQISGADGATFMTAPDT 760  
 QY 717 ITSVMANGFVISEDLGLITTPPELQAFOPPFILNLPYSVINGEPALEITTFNVLKD 776  
 Db 761 ITSMIFSFAVSVDHGLGVSEOH-KVTVFRNFPTLMLPVAVINGELIYQAIYFNLYST 819  
 QY 777 ATEVKVIEKDKPDILMTSSEINATGHQOTLLVPSEDAVTLFPIRPTHGELPIVTVA 836  
 Db 820 EVDVAVLTLESNKVTLARPGNNSAAGFSRITTPAGSSSVKPIRMGTIGELPIMTA 879  
 QY 837 LSPITASDAVTOMILVKAEGIEKYSQSILDLDTNRLOSTLTKLSFSPPTVTGSEBVQ 896  
 Db 880 ISEIASDALTRKVVQPEGITQCTSGSVLPQRMDASA.PPVESLNIOIPAGIVGSEKVK 939  
 QY 897 ITAIGDVLGSPINGLASILMRPYCGGEGONMNPANITYILDYTKKQOLDNLKEKLSF 956  
 Db 940 LLVVGDLIGSTMMNLGSLTRPSCCGEGONMGFAPDVFTLLYHSACKLADATRAKAFKH 999  
 QY 957 MRQGYORELYQORDGSFSAFGNVPDGSOTLSAFVLRCPF---EADPYIDIDONVHRT 1013  
 Db 1000 FQTSNSELNATKHRDGSFASGEGASGSTWLTLPAAFCFPAELRPTL-VSASVTDQA 1058  
 QY 1014 YTMKLGKQKSGEFGWDPGRVHSELQGNKSPVTLITAYIVSLGKRYKQPNIDVQSIH 1073  
 Db 1059 LTFILNQNTGTGTRRPGRVSHKAMQGVDSPIMTAVLITLKETNYAVANRAVQDAAE 1118  
 QY 1074 ----FLSEFSRGSISDN-YTLALITYALSSVSGSPKAEALMMLTWABQBGQMF--V 1125  
 Db 1119 NARIYENHILT-SISDNKYALAIYVALHVGSSRAEALLALBALATVQGGFKFMDNS 1177

QY 1126 SSESKLSDSWQ-----PRSLDIEVAAVALLSHFLQFQTEGPIPMKMLSRQNSLGGFAS 1180  
 Db 1178 ESPDSYSSRRRPYYNPPTNDIEMSAVALLTYRRDLNAGI.PVMKWLASKRSSLGGYSG 1237  
 QY 1181 TQDTYALKALSLSFALMNTERTNIQVYTYGSSPSPPL-----AVQPM--- 1224  
 Db 1238 TQDTYALIAQLSTYKAGLVLGNTQNLQISASHSDPPTASYNINRENSIVNSVNAVVDG 1297  
 QY 1225 AVNISANGFPAICOLNVVNVKASGSSRRRSIQNQ---EAFD-----LDVAVENKO 1275  
 Db 1298 TVQYATAGVAVAAQISVCNT-----PQOPIEFQCTNTVYSTRALKAK- 1344  
 QY 1276 DLNHVDLNVCTSF-SGPRGSGMALMEVNLGFMVP-SEASISLFTKRYKVEYDHGKMLTY 1333  
 Db 1345 -----VNMCCSLRPGDNATGMFLMEVNLPSGYTVAINDERTNPSAKLVEIDGNGVNVY 1398  
 QY 1334 LDSVNE-TQPCVNIIPAVRNFKVSNTODASYSIVDYIEPRQAVRSNSEVKLSSCDLCS- 1391  
 Db 1399 YDELAPGRSVCAIDIELINLGNVGSKARKVAADYQPKERVYALQVDEAPVVCDSCT 1458  
 QY 1392 -DVQGRPCED 1401  
 Db 1459 EDYAVCSVCAAD 1469

RESULT 7  
 081PH5  
 ID 081PH5 PRELIMINARY; PRT; 1388 AA.  
 AC 081PH5;  
 DT 01-MAR-2003 (TREMBlrel. 23, Created)  
 DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)  
 DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)  
 DE CG7052-PE.  
 GN TEPIT.  
 OS Drosophila melanogaster (Fruit fly).  
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 OC Ephydroidea; Drosophilidae; Drosophila.  
 NC NCBI\_TaxID=7227;  
 RX MEDLINE=20196006; PubMed=10731132;  
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,  
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
 RA Sutton G.G., Mortman J.R., Yandell M.D., Zhang Q., Chen L.X.,  
 RA Brandon R.C., Rogers Y.H., Blazer R.G., Champe M., Pfeiffer B.D.,  
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Gabor G.L.,  
 RA April J.F., Agbayani A., An H.J., Andrews-Pfannkoch C., Baldwin D.,  
 RA Ballwey R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,  
 RA Botkova D., Botchan M.R., Bouck J., Brokstein P., Brotier P.,  
 RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,  
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,  
 RA de Paaloe B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,  
 RA Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glaeser K.,  
 RA Glisler A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
 RA Hartsis N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,  
 RA Hostin D., Houston K.A., Howland T.J., Wei M.H., Ibegwam C.,  
 RA Jolali M., Kalush F., Kappen G.H., Ke Z., Kenison J.A., Ketchum K.A.,  
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
 RA Lasok P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
 RA Liu X., Matei B., McIntosh T.C., McLeod M.P., McPherson D.,  
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,  
 RA Mount S.M., Moy M., Murphy B., Murphy C., Murty D.M., Nelson D.L.,  
 RA Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Pacieb J.M.,  
 RA Palazzolo M., Peltman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
 RA Rainert K., Remington K., Saunders R.D., Scheeler F., Shen H.,  
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,  
 RA Slier E., Spradling A.C., Stapleton M., Strong R., Sun E.,

RA Svirskae R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
Wang Z.Y., Wasserman D.A., Weinstein G.M., Weisenbach J.,  
RA Williams S.M., Woodgett, Morley K.C., Wu D., Yang S., Yao Q.A., Ye J.,  
RA Zhen R.F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
RA Zhen X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.,  
RT "The genome sequence of Drosophila melanogaster";  
RL science 287:2185-2195 (2000).  
[12]  
RP SEQUENCE FROM N.A.  
RA Celinker S.E., Adams M.D., Krommiller B., Man K.H., Holt R.A.,  
RA Evans C.A., Gocayne J.D., Amanatides F.G., Brannon R.C., Rogers Y.,  
RA Banton J., An H., Baldwin D., Banton J., Beeson K.Y., Buesam D.A.,  
RA Carlson J.W., Center A., Champe M., Davenport L.B., Dietz S.M.,  
RA Dodson K., Dorssett V., Doup L.E., Doyle C., Drenek D., Farfan D.,  
RA Ferreira S., Frise E., Galle R.F., Garg N.S., George R.A.,  
RA Gonzalez M., Houch J., Hoskins R.A., Hostin D., Howland T.J.,  
RA Ibegwam C., Jalali M., Kruse D., Li P., Mattei B., Moshirei A.,  
RA McIntosh T.C., Moy M., Murphy B., Nelson C., Nelson K.A., Nuno J.,  
RA Pacleb J., Paragias V., Park S., Patel S., Pfeiffer B.,  
RA Phouanavong S., Pitterman G.S., Puri V., Richards S., Scheeler F.,  
RA Stapleton M., Strong R., Svirskae R., Tector C., Tyler D.,  
RA Williams S.M., Zaveri J.S., Smith H.O., Venter J.C., Rubin G.M.;  
RT "Sequencing of Drosophila melanogaster genome";  
RL Submitted (MAR-2000) to the EMBL/Genbank/DBJ databases.  
[13]  
RP SEQUENCE FROM N.A.  
RA Misra S., Crosby M.A., Matthews B.B., Bayraktaroglu L., Campbell K.,  
RA Hradecky P., Huang Y., Kaminck J.S., Prochuk S.E., Smith C.D.,  
RA Tupy J.L., Bergman C., Bertram B., Carlson J.W., Celinker S.E.,  
RA Clamp M., Drysdale R., Emmert D., Frise E., de Grey A., Harris N.,  
RA Krommiller B., Marshall B., Millburn G., Richter J., Russo S.,  
RA Seale S.M.J., Smith E., Shu S., Smutniak F., Whitfield E.,  
RA Ashburner M., Gelbart W.M., Rubin G.M., Mungall C.J., Lewis S.E.;  
RT "Annotation of Drosophila melanogaster genome";  
RL Submitted (MAR-2000) to the EMBL/Genbank/DBJ databases.  
[14]  
RP SEQUENCE FROM N.A.  
RA Adams M.D., Celinker S.E., Gibbs R.A., Rubin G.M., Venter J.C.;  
RL Submitted (MAR-2000) to the EMBL/Genbank/DBJ databases.  
[15]  
RP SEQUENCE FROM N.A.  
RA FlyBase;  
RL Submitted (SEP-2002) to the EMBL/Genbank/DBJ databases.  
DR EMBL; AB003618; AALN0639.1; -;  
SQ SEQUENCE 1388 AA; 154276 MW; FE57D37D7DDCE1E CRC64;  
Query Match 23.2%; Score 1703; DB 5; Length 1388;  
Beat Local Similarity 31.2%; Pred. No. 7e-95;  
Matches 458; Conservative 288; Mismatches 539; Indels 182; Gaps 50;  
QY 19 ALAAVAPGPFVLTAGIIRPGG--NYTIGVELLEHCPGQVYKAEILKTASULTVSVLEA 76  
DB 14 ALLVATAGTISVVGPGTLEKSNKYNVAVVHRADG--PSQ--IKVSLNGSPSYMET----- 64  
QY 77 EGVFEKSGFKTLTLPGLPLNSADEIYEL-----RVNGRTODELLENSRSLSPETKR 128  
DB 65 -----KQIELEP--PSTQWVEFEVYKATGATGNNISAEVSGVFKASKTKLMYADK 113  
QY 129 ISVFTQTDKALYKPKQEVKRIIVTLFSDPKPYKTS--LNLILKDPKSNLIQOMLSQOSDL 186  
DB 114 PSVFOVTDKATYKPADLVQFRILFDENRPAKIEKPIVISIIIDGQNRKIKQLSDVPLTK 173  
QY 167 GYISTKPOLSHPIILGMSIQOVN--DQYTYOSFOVSEVYLPKFEVTLQTLPLCYSMNSKH 245  
DB 174 GVFSESELQISEQPVIGTWKISVSDGNNKETSFEVDKYLPLKFEIYIVDTPRAVVIADKY 233  
QY 246 LAGTITAKTYGKPKVKGDTLTF-----FLPLSFVKGKGNITKPKFKINGSANFSNDE 298  
DB 234 IKATIRAKTYGKPKVKGDTLTF-----FLPLSFVKGKGNITKPKFKINGSANFSNDE 288  
QY 299 MKNVDDSSNGLEBYLDLSSPGVEILITVTESVTGISRNVSSTNVFEPKQDIYIEFDYTT 358

DB 289 --IIHMAQRG--QYLP-----PIKLFVAVTEELTGKNNKATATVVLVHQORYSIEPERE 339  
QY 359 VLKPSLNTFATVATKVRADGNOLTEERRNNVITVQRYNTYEWGSGNSGKQKEAVOKI 418  
DB 340 HFPAKNSITTYOVVKNVNDGSPVT--NSAKVNIKGF--DKSYVFFHPSPK-----TRI 388  
QY 419 NYTVPOQS-----GTFKIEPILDESSELOKAYFLGSKSMAVHSLFK--SP--SKTYIQDKT 472  
DB 389 NFEAPVNEGLATEFVRLPD--SDSRYYRIASFPOSENTISISIKFEETPLMSPREPLKIQV 447  
QY 473 RDENIKVSPPELVGSKKRLKELSLMYVSGQL-----VANGKNSYTFSLTPNSW 525  
DB 448 NTKKRIEQQVSPDVSIEDLPFYFYITIVAGANVILSDYVDVPPQOKYTYV--KFTPFPM 506  
QY 526 TPKACVITYYIEDGEIISDVLEKI PVOLFNGKIKLWYKSKVAKPSEKVSIRISTQDPS 585  
DB 507 VPKATITYYVAVN--DIOFEKRTIDFEKFSNSIDVS--APINAKPSEEVKLRK--TDADS 563  
QY 586 IVGIYAVADKSVNLMAASNDITMENVHLELYNTGYILGFMNSFAV--QECGLVLTLD 643  
DB 564 FVGLIGVDQSVLLKSGNDLSQDDI FNSLNTYQTS---TPMNGYGRYPGQTSGLVLTIN 620  
QY 644 ANLTQYIDGV--YDNAEYAEKFMENBSGHVYDHFSLGSSPH--YKXHPETWYLDLT 699  
DB 621 ANPYNTEDSISYD-----EVDAISITSSTKLELVRTNFAEYVMMTTS 663  
QY 700 NMGRYIVQEFVNT--VPDSITSWATGPFVISEDGLGTTTPVLEQAFQPFIFLNPYS 757  
DB 664 DNGSVGESEFVLTTKIPITITSWVATGSLPFTSGIALTKPKSKIRVQPPFVSTNLNLS 723  
QY 758 VIRGEFALTEITFNYLKDQATEVAVILKSD--KEDILMTSSE-----INATGHOQTLNLP 811  
DB 724 VKRGVIALPVIYFNYLTKTDADAVMNSQOEYEFTEATNEVLEKALIDEVRVAVRTIP 783  
QY 812 SEDGATVLPPIRPHLGIPIPTVLTALSPASDAVOTMLVKAEGTISKYSQSLIDLTQN 871  
DB 784 ANSGKSVSEFMRPKVNGVFTTKITATSLADDAIHQKLKVEPEGVTLFENRAVFINLADQ 843  
QY 872 RLQSTLKLTSFPFNNTYTGSEVQITAGVNLGSLNGSLRMFPGCEQNNINAP 931  
DB 844 PEMS--QSLDADIPREVPOSEFIEFSVAGDLPELTQNLNLVMPGCEQNNVNVVP 901  
QY 932 NIYILDYLTKKKQTLNKEKALSFMRQGYORELLYQREDSFSFAGVYDPSGTWLSAF 991  
DB 902 NILVILKYLEVGRKLPSTBSAKKFLBELGYRELTAKHDDSYSAFGSDASGSTMWLTAY 961  
QY 992 VLRGFLADPYIIDQVYARTTYTWLKGHQKSGNEFMDPGFVHISELOGKNSPYTLTAY 1051  
DB 962 VMRSFHQGTGTVDIDPKVITAGLDPLVSKQESGEFPEVGL-----FDNANQNPLALTSF 1017  
QY 1052 IYTSILGTRKQPNID--VQESIHFLSEFSGISDNLTALITYALSSVSPKAKELN 1109  
DB 1018 VLAFFENHELDIPKYQSAIKKAVYVAEADK--TDQVSLIAVAVALQALAHGPOSEKYIA 1076  
QY 1110 MLTWRAEQQGKQFQVSSSEKSLDS-----WQPSLIDIEVAAYVALSHPLQFQSE 1160  
DB 1077 KLEYSARKENBRMW-----STATESGEGDVHFKKRSNDVELITSYVLLLEKDPREK 1132  
QY 1161 GIPIMWLSRQNSLGFASQTDTTVALKALSEFA--ALMNTERTINQVTVTGPS----- 1213  
DB 1133 ALPIIKWILISQRNSNGSGSTQDPTVIGLALTFAFKYTGSGSGTMDIFSSAGESKNTIK 1192  
QY 1214 -SPSPILAIVQ-----PMANVISANGGFAICQLNVYVYNYKASGSSRRRSIQOEAFD 1265  
DB 1193 VNPENSLVLQTHDLPEKSTRKVDFTAKGSAVMDLSRYVYNAEK-----EKKPSFK 1243  
QY 1266 LDVAVKEKDDLDNVDLNVCTSF-----SGPGRSGMALMEVNLISGFMVSEA---LSL 1316  
DB 1244 VTPYVQDIPNQLLVD--VCAEYVPLEBADXDQXSNMAVMIALPSGCVGSTISGKIQ 1301  
QY 1317 SEYVAKVEYDH--GELNLVYDSVNETQF--CVNIPAVNFVKSNTQDASVSIVDYIEPRRQ 1373  
DB 1302 VDRVKEVETKNSDSITVVYFDSLTPGDVRCPLPASKAHAAVAKQKPAVASLYDYDTERK 1361

1374 AVRSYSEVKSCLDCLSDV---OGCR 1397  
1362 ATEY--QVSSLDICGAGCGGCK 1386

## RESULT 8

09VL20 PRELIMINARY; PRT, 1397 AA.  
01-MAY-2000 (TREMblrel. 13, Created)  
01-OCT-2002 (TREMblrel. 22, Last sequence update)  
01-MAR-2003 (TREMblrel. 23, Last annotation update)  
CG7052 protein.  
TEP1 OR CG7052 OR CG18589.  
Drosophila melanogaster (fruit fly).  
Eukaryota, Metazoa, Arthropoda, Hexapoda, Insecta, Pterygota;  
Neoptera, Endopterygota, Diptera, Brachycera, Muscomorpha;  
Ephydroidea, Drosophilidae, Drosophila.  
NCBI\_TaxID=7227;  
[1]  
SEQUENCE FROM N.A.  
RC STRAIN=Berkeley;  
RA MEDLINE=20196006; PubMed=10731132;  
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,  
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,  
RA Brandon R.C., Rogers Y.-H.C., Blazer V., Champs M., Pfeiffer B.D.,  
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,  
RA Abail J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,  
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,  
RA Botkova D., Botchan M.R., Bouck J., Brockstein P., Brothier P.,  
RA Butts K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,  
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,  
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,  
RA Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,  
RA Glander A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
RA Haddix N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,  
RA Hostalka D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,  
RA Jaitani M., Kalush F., Karen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
RA Laeko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,  
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,  
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
RA Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Paclab J.M.,  
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,  
RA Spier E., Spradling A.C., Stapleton M., Strang R., Sun E.,  
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,  
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao X., Zheng L.,  
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
RT "The genome sequence of Drosophila melanogaster";  
RL Science 287:2185-2195(2000).  
RN [2]  
RP SEQUENCE FROM N.A.  
RA Celniker S.E., Adams M.D., Kronmiller B., Wan K.H., Holt R.A.,  
RA Evans C.A., Gocayne J.D., Amanatides P.G., Brandon R.C., Rogers Y.,  
RA Batzon J., An H., Baldwin D., Batzon J., Beeson K.Y., Busam D.A.,  
RA Carlson J.W., Center A., Champs M., Davenport L.B., Dietz S.M.,  
RA Dodson K., Dorsett V., Doup L.E., Doyle C., Dresnek D., Fattan D.,  
RA Ferreria S., Frise E., Galie R.F., Garg N.S., George R.A.,  
RA Gonzalez M., Houck J., Hoskins R.A., Hostalka D., Howland T.J.,  
RA Ibegwam C., Jaitani M., Kruse D., Li P., Mattei B., Moshrefi A.,  
RA McIntosh T.C., Moy M., Murphy B., Nelson C., Nelson K.A., Nuno J.,

PA Paclab J., Paragas V., Park S., Patel S., Pfeiffer B.,  
PA Pounanavong S., Pittman G.S., Puri V., Richards S., Scheeler F.,  
PA Stapleton M., Strong R., Svirskas R., Tector C., Tyler D.,  
PA Williams S.M., Zaveri J.S., Smith H.O., Venter J.C., Rubin G.M.,  
RT "Sequencing of Drosophila melanogaster genome";  
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
RN [3]  
RP SEQUENCE FROM N.A.  
RA Miera S., Crosby M.A., Matthews B.B., Bayraktaroglu L., Campbell K.,  
RA Hradecky P., Huang Y., Kaminker J.S., Prochuk S.E., Smith C.D.,  
RA Tupy J.L., Bergman C., Berman B., Carlson J.W., Celniker S.E.,  
RA Clamp W., Drysdale R., Emmert D., Frise E., de Grey A., Harris N.,  
RA Kronmiller B., Marshall B., Milburn G., Richter J., Russo S.,  
RA Seale S.M.J., Smith E., Shu S., Smutnack F., Whitfield E.,  
RA Ashburner M., Gelbart W.M., Rubin G.M., Mungall C.J., Lewis S.E.,  
RT "Annotation of Drosophila melanogaster genome";  
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
RN [4]  
RP SEQUENCE FROM N.A.  
RA Adams M.D., Celniker S.E., Gibbs R.A., Rubin G.M., Venter C.J.;  
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
RN [5]  
RP SEQUENCE FROM N.A.  
RA FlyBase;  
RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AEO03618; AAF52540.2; -;  
DR HSSP; P01024; 1C3D.  
DR FlyBase; FBgn0041182; Tep1.  
DR InterPro; IPR002890; A2M\_N.  
DR InterPro; IPR001599; MacrolobinA2.  
DR InterPro; IPR000215; Serpin.  
DR Pfam; PF00207; A2M; 1.  
DR Pfam; PF01835; A2M\_N; 1.  
DR PROSITE; PS00477; ALPHA\_2\_MACROGLOBULIN; 1.  
DR PROSITE; PS00284; SERPIN; 1.  
SQ SEQUENCE 1397 AA; 155108 MW; 9CE9B684AB207574 CRC64;

Query Match 23.1%; Score 1699.5; DB 5; Length 1397;  
Best Local Similarity 31.3%; Pred. No. 1.2e-94;  
Matches 460; Conservative 288; Mismatches 543; Indels 179; Gaps 52;

19 ALAVAGPREPLVTPGIIIRPG--NVTIGVELHCPSQVTVKAEIKTASNLTVSLA 76  
14 ALVYMTGIVSVGPGPTLSNSKYNVVSVHKADG--PSQ--IKVSLNGSYNET----- 64  
77 EGVEFKSGFRTLLPLBPLNSADEIYL-----RVGRTQDELFSNSTLSPETKR 128  
65 -----KQIELP--PMTQVVEVEVPKATGNTNLSHGVSQVFNKSTKLNTADKK 113  
129 ISVFIOTDKALYPRKQEVKRYITLTFSDFKPYKTS--LNTLIDPKSNLIQWLISQOSDL 186  
114 PSFVFGTQDKATYPRALVGRILFLDENRPAIEKIPISVIIIDGQNRKQISDVKLTK 173  
187 GVISKTFQSLSHPLIDMSIQOVN--DQTYGQFQVSEYLLPFPEVTLQTLPLCSMNSKA 245  
174 GVPSGSELQISEQVGLTGWKISVSDVDNRETGKFEVDKVLPRFEYIVDTPKAVVADKV 233  
246 LAGTITAKTYTGKPVGVDTLT-----FLPISFGKKKNTIKTKINSANFSFINDER 298  
234 IKATIRAKTYTGKPVGVKATVSMERSYGFGLDANANGNQE--KTIDVGKGHVEPD--- 288  
299 MKVMDSSNGLSEYLLSPGPEILITVTESVTSISRVNVTVPFKQHYIIEFDYTT 358  
289 --IIHWAQRG--QYLP-----PIKLAVTTELTGKQKATATVVLHQQRYSIEPERPE 339  
359 VLKPSINFATYKVTADGNQLTLERRNNVITVQRTYTEVSGSNGNQMEAVQKT 418  
340 HFAKNSFTYQVVAVKVDGSPVT--NSAKVVKIGF--DKSYSYHEBSPK-----TRI 388  
419 NTVVPGS---GTFKIEPFLDESSSELQKAVPLGSKSSNAVHSLK--SP--SKTYIQLKT 472  
389 NFAPAVNENGIATFNRLPD--SDSRYYRIFAPSGDSENTIGSISKEPTMSREPLKIQV 447

QY 473 RDNENKVSPEFLVYSGNKLKELSYMVVSRGOL-----VAVGKONSTMSLTPENSW 525  
 Db 448 NTKKCRRLGBOVSFDVVSIEDLPFYFYITVARGNVILSDVVDVDPDQKITYTV-KFTPTESM 506  
 QY 526 TPKACVIVYIEDDBEIIISDVAKIPQVLFKKIKLYMSKVAESEKVSILSTYQDS 585  
 Db 507 VPKATVYVYVNN-DLOPEEKTIQPEKEFSNIDVS-APINAKSEEVKILIK-TDADS 563  
 QY 586 IVGIYAVDKSVLNNASNDITWENVHELELYNTGYGLGMPNNSPAVF--QEGCLMVLTD 643  
 Db 564 FVGLIGVQSVLLKSGNDLSGDDLFNSLNITYOTS---TPMANGIGRYGQISGLVTLTN 620  
 QY 644 ANLTKOYIDGVYDVAERF-MEENEGHIYDIDHDSL-----GSSPHVKHPPETWTL 697  
 Db 621 ANYPY-----NTEPPIAFSLAQAIAAGMPGSSIASHPNOAPQIRKEFPENWIFY 672  
 QY 698 DT-MNGYRIYQEFYV--VPDSITSWATGFIYSEDLGGLITTTVELQAPPPFIPLNL 754  
 Db 673 NAEVNG---EEEFPTLKIPDITISWVYVGFSLNPTSGIALTKNSPKIRVFPFVSTNL 729  
 QY 755 PYSVIRGEFELEITIFNYLKDTEVKVIEKSD-KFDILMTSS-----INATGHQOOL 808  
 Db 730 PYSVIRGEVIALPVYIFNYLKDTEVDVMDNSDDEYETETEXTNVELEKALDEVRVAKV 789  
 QY 809 LVPSDGAIVLFPPIRPHLGEIPITVYALSPTASDAVTOMILVKAEGIEKYSQSLIDL 868  
 Db 790 TIPANSKGSVSEFMIPKRVYFTLLKITATLALAGAHQKLVBEQEGVTLFENRAVFINL 849  
 QY 869 TDNRLOSTLKTSPFPRTVYSGSERVOITAGDVLGSPINGLASIRPPYCGGQNMN 928  
 Db 850 KDQPEMS--QSLDADIPEWVQSEFIEBSVVDLGLPTLOMLVVRNPYCGGQNMN 907  
 QY 929 FAPNIYIIDLKQKQOLTDNLKEKALSFMRQGYORELYOREDGSAFAGNVPDSQWTL 988  
 Db 908 FVFNILVAKILEVYTGKRLPSVSKARKLEIGYQELTYKHPDGSYSAFSGDSAGSTWL 967  
 QY 989 SAFVLRCELEADPYIDIDQVLYARTYMLKQHOKSNGEWDPRVYHSELQGNKSPVYL 1048  
 Db 968 TAYVRSFHQAGTYDIDPQKITAGDLPLVSKQSGSEPEVQKL---FDNANQPLAL 1023  
 QY 1049 TAYVTSILGRKTYQPNID--VQESIHPLESEFSRSGISDNTYLLITLALSVSQPKAKE 1106  
 Db 1024 TSFVLAFFENHELLPKYQSAIKKAVRYAEADK-TDOYSLATAVAVALQAKHPOSEK 1082  
 QY 1107 ALNMLTWAEBOGQWVSSSEKLSDS-----NOPRDLDEVAAYALSHFLOQ 1157  
 Db 1083 VIAKLESAVARKENDRMM---SKATESTGEDGRVFMKPRSDVEITSYVLLALLEKDP 1138  
 QY 1158 TSEGIPIRNLWLSRQNSLGGFASITODTVALKALSEFA--ALMTERTNVQVTVGPS-- 1213  
 Db 1139 AEKALPIIKMLISQNSNGSFSTODTVIGLOALTKFAVKGSGSGTMDIESSAGESEK 1198  
 QY 1214 ----SPSPPLAVQ-----PMVANIISANGFPAICQILNVVYNKASGSSRRRSIQOE 1262  
 Db 1199 TIKVNPENSILVQTHDLPEKSTRKVDFTAKGTGSAMVQSYRYNLAEK-----EKKP 1249  
 QY 1263 AFDLDVAKENKDLNHDVLANCTSF-----SGGRSGMAMENVNLISGVNPSA--- 1313  
 Db 1250 SFKVPTVADFPNQLIYD--VCAAYVPLEADAKDKKDSMAVAMELALPBGFGDSTSLGK 1307  
 QY 1314 ISLSTVKKVEYDH--GKILNLYDSVNETOF--CVNIPAVRNKFNVTODASVYIDYEP 1370  
 Db 1308 IQAVDVRKVRKETSNDSTVYVFDLSLTPDVRCLPLEASKAAVAVAKQKRPASVSLDYDT 1367  
 QY 1371 RRQAVRSNSSEVKSJSCDSDV---QGR 1397  
 Db 1368 BRKATEYY--QVSSSLCIDCEGADGCGCK 1395

RESULT 9  
 Q81PH4 PRELIMINARY; PRT; 1408 AA.  
 AC Q81PH4;

DT 01-MAR-2003 (TrEMBLrel. 23, Created)  
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)  
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)  
 DE CG7052-PD.  
 GN TEPII.  
 OS *Drosophila melanogaster* (Fruit fly).  
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 OC Ephydroidea; Drosophilidae; *Drosophila*.  
 OC NCBI\_Taxid=7227;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=20196006; PubMed=10731132;  
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,  
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
 RA Sutton G.G., Wortman J.R., Vandeil M.D., Zhang Q., Chen L.X.,  
 RA Brandon R.C., Rogers Y.H., Blazey R.G., Champe M., Pfeiffer B.D.,  
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Gaber G.L.,  
 RA Abril J.F., Agbayani A., An H.J., Andrews-Pfannkoch C., Baldwin D.,  
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,  
 RA Borkova D., Botchan M.R., Bouck J., Brostein P., Brotlier P.,  
 RA Burris K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,  
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,  
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferrieres S., Fleischmann W.,  
 RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glaser K.,  
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,  
 RA Hoston D., Houston K.A., Howland T.J., Wei M.H., Ibegwam C.,  
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Kerchum K.A.,  
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
 RA Lasro P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,  
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,  
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
 RA Nelson D.R., Nelson K.A., Nixon S., Pollard J., Puri V., Reese M.G.,  
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
 RA Reinert K., Remington K., Saunders R.D., Scheefer F., Shen H.,  
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,  
 RA Splitter B., Spredling A.C., Stapleton M., Strong R., Sun E.,  
 RA Svrtkals R., Tector C., Turner E., Venter E., Wang A.H., Wang X.,  
 RA Wang Z.Y., Wasserman D.A., Weinstein G.M., Weissbach J.,  
 RA Williams S.M., Woodger J., Worley K.C., Wu D., Yang S., Yao Q.A., Ye J.,  
 RA Yeh R.F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
 RT "The genome sequence of *Drosophila melanogaster*.";  
 RL Science 287:2185-2195 (2000).  
 [2]  
 RN SEQUENCE FROM N.A.  
 RP Celniker S.E., Adams M.D., Kronmiller B., Wan K.H., Holt R.A.,  
 RA Evans C.A., Gocayne J.D., Amanatides P.G., Brandon R.C., Rogers Y.,  
 RA Barton J., An H., Baldwin D., Barton J., Beeson K.Y., Busam D.A.,  
 RA Carlson J.W., Center A., Champe M., Davenport L.B., Dietz S.M.,  
 RA Dodson K., Dorselt V., Doup L.E., Doyle C., Driesnek D., Farfan D.,  
 RA Ferrieres S., Friese B., Galle R.F., Garg N.S., George R.A.,  
 RA Goncalves M., Houck J., Hoskins R.A., Hoston D., Howland T.J.,  
 RA Ibegwam C., Jalali M., Kruse D., Li P., Mattei B., Moshrefi A.,  
 RA McIntosh T.C., Moy M., Murphy B., Murphy L., Nelson K.A., Nunoo J.,  
 RA Pacleb J., Paragas V., Park S., Patel S., Pfeiffer B., Pfeiffer F.,  
 RA Phouanavong S., Pittman G.S., Puri V., Richards S., Scheefer F.,  
 RA Stapleton M., Strong R., Svrtkals R., Tector C., Tyler D.,  
 RA Williams S.M., Zaveri J.S., Smith H.O., Venter J.C., Rubin G.M.;  
 RT "Sequencing of *Drosophila melanogaster* genome.";  
 RL Submitted (MAR-2000) to the EMBL/Genbank/DBJ databases.







D	b	564	FVGLLGVQSVLLTKSGNDLSQDDIFNSLINVQTS---TPMNVGGRVQGQSGVLTLLN	620
Q	y	644	AN-----LKQDY-----DGVONAEAEAFMBE--NEGIIVIDIHDS	679
D	b	621	ANYPNTGPLVMSVYFEGSRHFWITPRYRVGIRDDSGDRISFLQSLDRNLKEI---L	677
Q	y	680	LGSSEH---VRKHPEPETYIMLDTNMGYRIYQFEVY--VPDSITSMWAGFISEDLIG	734
D	b	678	LKQIPQRTTIKEFEETWFF--ENVG---EEFFLITLKIPDITISMVYVGFSLNPTSGIA	732
Q	y	735	LTTPPELQAFQPPFIPLNLPYSVIRGEEFALEITFENYLKDATEKVIIEKSD-KFDIL	793
D	b	733	LTQPSKIRVQPFQFVSVTNLPYSVRGGEVIAIPVIVFNULDITLADVMDNSDDEYEFT	792
Q	y	794	MTSS-----INATGHOOITLVPSDEGATVLEPIRPHLGEIPIYVTLASPASAVQM	848
D	b	793	EATNEVELEKAIDVRRKRVKVTIIPANSGSVSMIRPKVNGFTLLIKITATISALAGDIHOK	852
Q	y	849	ILVYAEIGIEKSVQSILLDITDNRLOSTLKTLSFSFPNTVYSGSERVOITAGDVLGSI	908
D	b	853	LKVEPEGVTLFENRAVFINLKQOPEMS--QSLDADIPEVVPQSEFIERSVGDLGLPTL	910
Q	y	909	NGLASLIMFPGCCGGMNINFPNVIYIDYLTKKQOLDNLKEKALSPKROGYQRELYQ	968
D	b	911	QNLNLMVMPYGCQGMNVNFPNLTIVLYKULEVTRKJPSVBSKARKFELIYGRELYTK	970
Q	y	969	REDSFSAFGVNDPEGSGTMSAFVRCFLEADPYIDIDONVLAHRYTYMLKQKNGEPM	1028
D	b	971	HDDGSYSAFGSDAGSGTMLTVYVRSFHQAGTYDIDBKVTIAGDLPVLSQKSGEFP	1033
Q	y	1029	DGPRVHSELQGGNKSPTVLTAYIVTSLGKRYKQPNID--VOESIHPLESEFSHGISDN	1086
D	b	1031	EVGKI-----PNANONPNPLATISFVLAAFFENHELIPKYOAIKKKAVRYAEBADK-TDQ	1085
Q	y	1087	YTLALITYALSSVSGPKAKEALNMLTWBAEOHCGMOPWVSSSKXISDS-----MOP	1133
D	b	1086	YSLAAVALQAKHPQSEKVIAKLESYARKENDMMW---SKATSESGEDGRVFWHXP	1144
Q	y	1138	RSLDIEVAAYALLSHFLQFOTSEGIPIKMWLSRQNSLGSFASOTDITYALKALSEFA-	1195
D	b	1142	RSNDVEITSYVLLALLEKROPAEKALPIIKWILISQNSNGSFGSTQDVTYIGALTKPAVK	1200
Q	y	1196	ALAMTERINIOVTVTGP-----SPSEPLAVQ-----PMAVNIISANGFPAICQNAV	1244
D	b	1202	TGSGSGTWDIEFSSSAGESKNITIKVPENSLVLOTJDLPKSTRKVDFTAKGTGSANVOLSY	1263
Q	y	1243	VYNYKASGSSRRRRISIQONCEAFDLVAIVKENCDLNVHVDLVNCTSF-----SGPGRGM	1296
D	b	1262	RYNLAEK-----EKKPSFKVPIYKQDIPNQLLIYD--VCAEYVPLEADKXKDSNM	1311
Q	y	1297	ALMEVNLISGFMPVPSEA---ISLSEYTKVEYDH--GKLNLYLDSVNETQF-CVNIIPVR	1350
D	b	1311	AVMEIATPSGFVGDSITSGKIQADVVRKRVETKNDSITVVYVFDSLTPGDVCLPLEASK	1370
Q	y	1351	NPKYSNTODASVSIYDVYEPERRQAVRSYNSYETSSCOLGSV---QQCR	1397
D	b	1371	AHAVAKQPAVASLYDYVDEKAEYV--QYKSLCOTICGADCGEGCK	1418

[1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Berkeley1  
RC MEDLINE=2019606; PubMed=10731132;  
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,  
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
RA Sutton G.G., Mortman J.R., Yandell M.D., Zhang Q., Chen L.X.,  
RA Brandon R.C., Rogers Y.-H.C., Blazey R.G., Chagne M., Pfeiffer B.D.,  
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,  
RA Abail J.F., Agbayan A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,  
RA Ballew R.F., Baugu A., Baxendale A., Baxendale J., Bayraktaroglu L., Beasley B.M.,  
RA Beeson K.Y., Benos P.V., Bertman B.P., Bhandari D., Bolshakov S.,  
RA Borokova D., Botchan M.A., Butler H., Cadieu E., Center A., Chandra I.,  
RA Butts K.C., Buam D.A., Butler J., Cadiu E., Center A., Chandra I.,  
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,  
RA de Paolis B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,  
RA Foster C., Gabrielian A.E., Gang N.S., Gelbart W.M., Glasser K.,  
RA Glodex A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,  
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibeagwam A.,  
RA Hotell M., Kalush P., Kapen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
RA Kammel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
RA Lako P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
RA Lin X., Matei B., McIntosh T.C., McLeod M.P., McPherson D.,  
RA Mekulov G., Mishina N.V., Mobarry C., Morris J., Moshrefi A.,  
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Paclab J.M.,  
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,  
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
RA Wang Z.-X., Wasserman D.A., Weinstock G.M., Weissbach J.,  
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith L.,  
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.,  
RL "The genome sequence of *Drosophila melanogaster*." ;  
RL Science 287:2185-2195(2000).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC Celniker S.E., Adams M.D., Kronmiller B., Wan K.H., Holt R.A.,  
RA Evans C.A., Gocayne J.D., Amanatides P.G., Brandon R.C., Rogers Y.,  
RA Carlson J., An H., Baldwin D., Banzon J., Beeson K.Y., Busam D.A.,  
RA Barton J.W., Center A., Chagne M., Davenport L.B., Dietz S.M.,  
RA Dodson K., Dorett V., Doup L.E., Doyle C., Drenth D., Fartin D.,  
RA Ferreira S., Frise E., Galle R.F., Gang N.S., George R.A.,  
RA Gonzalez M., Houck J., Hoskins R.A., Hostin D., Howland T.J.,  
RA Ibeagwam C., Jaitai M., Kruse D., Li P., Matei B., Moshrefi A.,  
RA Mcintosh T.C., Moy M., Murphy B., Nelson C., Nelson K.A., Nuno J.,  
RA Paclab J., Paragas V., Park S., Patel S., Pfeiffer B.,  
RA Phouenavong S., Pittman G.S., Puri V., Richards S., Scheeler F.,  
RA Stapleton M., Strong R., Svirskas R., Tector C., Tyler D.,  
RA Williams S.M., Zaveri J.S., Smith H.O., Venter J.C., Rubin G.M.,  
RT "Sequencing of *Drosophila melanogaster* genome." ;  
RL Submitted (MAR-2000) to the EMBL/genbank/DBJ databases.  
RN [3]  
RP SEQUENCE FROM N.A.  
RA Maier S., Crosby M.A., Matthews B.B., Bayraktaroglu L., Campbell K.,  
RA Hradecky P., Huang Y., Kaminker J.S., Prochnik S.E., Smith C.D.,  
RA Tupy J.L., Bergman C., Bertram D., Carlisle J.W., Celniker S.E.,  
RA Clamp M., Drysdale R., Emmert D., Fritze B., de Grey A., Harris N.,  
RA Kronmiller B., Marshall B., Millburn G., Richter J., Russo S.,  
RA Seearle S.M.J., Smith E., Shu S., Smutnicka F., Whitfield E.,  
RA Ashburner M., Gelbart W.M., Rubin G.M., Mungall C.J., Lewis S.E.,  
RT "Annotation of *Drosophila melanogaster* genome." ;  
RL Submitted (MAR-2000) to the EMBL/genbank/DBJ databases.  
RN [4]  
RP SEQUENCE FROM N.A.  
RA Adams M.D., Celniker S.E., Gibbs R.A., Rubin G.M., Venter C.J.,

RL Submitted (MAR-2000) to the EMBL/Genbank/DBJ databases.  
RN [5]  
RP SEQUENCE FROM N.A.  
RA Flybase;  
RL Submitted (SEP-2002) to the EMBL/Genbank/DBJ databases.  
DR EMBL; AE003618; AAF52539.2; -.  
DR HSSB; P01024; IC3D.  
DR Flybase; FBgn0041182; Tepit.  
DR InterPro; IPR002890; A2M.N.  
DR InterPro; IPR001599; MacroglloblnA2.  
DR InterPro; IPR000215; Serpin.  
DR Pfam; PR00207; A2M; 1.  
DR Pfam; PF01835; A2M.N; 1.  
DR PROSITE; PS00477; ALPHA\_2\_MACROGLOBULIN; 1.  
DR PROSITE; PS00284; SERPIN; 1.  
SQ SEQUENCE 1399 AA; 155783 MW; 3D8B7DF91E39F59 CRC64;

Query Match 23.0%; Score 1688.5; DB 5; Length 1399;  
Best Local Similarity 31.3%; Pred. No. 5.4e-94;  
Matches 461; Conservative 285; Mismatches 544; Indels 183; Gaps 52;

19 ALAVAGPRLVTAGIIRPGG--NVTIGVELLEHCPQVTKAKELTKASLTVSLA 76  
14 ALAVATGIVSVGEGTILSKSNKTNVAVSHKADG-PSQ--IKVSLNGPSYNET----- 64  
77 EGVFEGSEFKTLTLPSPILNLSADEIYEL-----RVGRTQDEILFSNSTRLSFETKR 128  
65 -----KQELP--PMSTQNEFEVPRKATNGVNLNLSAEVSGVFNSTKLVYADK 113  
129 ISVFQITDKALYKPKQEVKRIVTLPSDFKPYKTS--LNLILKDPKSNLIQMLSQSD 186  
114 PSVFQITDKATKPADLVQFRILFLDENTRPKIKPISVIIIDAQVRIRKQLSDVKLT 173  
167 GVISTFQSLSHPIIGDMSIOVQN--DQTYSGQVSEVLPKFEVTLQTPICYSMNGH 245  
174 GVFSEGLQISEQPVLTGKIVSVGDNRKETSFEVDKYLPRFEVITDTPAAVVIADKV 233  
246 LINGITAKTYGKPKVGDVTLT-----FLPLSEWKKKNTTKFKINGSNANFSFND 298  
224 IKATIRAKTYGKPKGKATVSMERSYGFQDLNNGKQD--KIIDVDGKHVDFD--- 288  
229 MKNVDSNGSLSEYLDLSSPGVEILTVTTSVGTISRNVSNTVFPKQHDYIIIEFDYTT 358  
289 --IIMHAGG--QYL P-----PIKFAVYTEELTGKQNAVTAVLHQORYSIEBERPE 339  
359 VLKPSLNFATYKTRADGNQLTLEBRNNVITYTQRYTYMGSGNSGNGKMAVQKI 418  
340 HFEAKSFYIYQVAVKNDGSPVT--NSAKNVKIGF-DKSYSYFHEBSPK-----TRI 388  
419 NTVFPOS-----GTFKIEPILDESSELQKAYFLGSKSMAVHSLFK-SP-SKTYIOLKT 472  
389 NFEAVNENGATITFNVRLPD-SDSKRYRIFASFDOSENTIGISKEFEPNPSREPLKIOV 447  
473 RDNIRKVGSPFELVSVGNKRLKELSYMVVSRQL-----YAVGKQNSTMSLTLPNSW 525  
448 NTKKRLPQGVQSFVDVSIEDLPYFYTYIARGNVILSDVVDPDQKTYTV-KFPTTSM 506  
526 TPKACVIYVYIDDEGIISDVLKIPQVLVFKKIKLYSKVKAASEKSLRISTYQDS 585  
507 VKKATYYVYVNVN-DLOFEKTIIFEKKFSNSIVS-APTAKKSEBEKRLIK-TDDDS 563  
586 IYGIIVAVDKSVLNMANSNDITMENVHELENTGYILGMFNNSPVPV--GECSGLAVLTD 643  
564 FVGLGVDSDVLLLSGNDLSODDIFNSLNTYQTS--TPMNGIGRYFGQISGLVTLTN 620  
644 ANLTQDYIDGVYDNEAERFMEENEGHIVDHD--FSIGS-----SPHYRKHFPEWT 694  
621 AN-----YPNYTERIFIRPGIGFPRPLFNRYTVAGSLPPNVIPRQYRKPEPM 671  
695 IM-LDTNMGKRIYQREVT--VPDSTSVNATGFISELGLITTPPELOAFQDFPFLF 751  
672 IENIFENVG---EEEFLLTKKIPDITTSVVVGFSLNPTSGIALTKNPSKIRVQPFVFS 728

QY 752 LNLPSYVIRGEFALLEITITFENYLKDATEVKYIIKSD-KEDILMTSSE-----INARGHQ 805  
DB 729 TNLPSYVRGGEVIAIPVYIFNYLDKTLDDVMDNSDQYEFTBATNEVLEKALIDEVARY 788  
QY 806 QTLVPSDDATVLPFRPHLGEIPITVNLSTPADAVQMLVYKAEIGEKYSQSL 865  
DB 789 KRVITPANSKGSVSPMRPKNVGFTTLKITATSLAGDAIHQKIKVEEGVTLFENRAVF 848  
QY 866 LDLTDNRLOSTLTKLTSFPPNVTGSEVVOITAGVLPISINGLASLIMPYGCGEQN 925  
DB 849 INLQDQPEMS--QSLDADIPHEVVPQSEFIEFSVGDLLGFTLQNLNLTVMFPICGCEQN 906  
QY 926 MINEPANIYIIDLTKKKQLTDNLKELASFMRQGYORELLYQREDSFSAFGVDPBGS 985  
DB 907 MNFEPNIVLVKYLEVTRGKLPVSVEKARKPELIGYORELTYKHDSDGSYAFGSDASGS 966  
QY 986 TWLSAFVIRCLLEADPYTIDQNVLHRTYTLKQHSNGSEFMPPRGVHSELQGNKSP 1045  
DB 967 TWLTAVYWRSPHQAGTYTIDDPKVITAGLDFLVSKQKSGEFPEVGRKL---FDNANQNP 1022  
QY 1046 VTLTAVYITSLGVRKYQPNID--VQSIHPLSEFSFGISDNYTLALITYALSVSQSPK 1103  
DB 1023 LALTSFVLLAFENHEHLIPKQSAIKAVRYVABADK-TDDQYSLIAVALLQAKHPQ 1081  
QY 1104 AKELNMLTWRAEAGEGQWQVWSSSEKLSDS-----WQPSRLDIEVAVALSHFL 1154  
DB 1082 SEKIYAKLESVARAKENDRMW---SKATESTGEDGRVPHMKPSPNDVEITSYVLLALLE 1137  
QY 1155 QFOTSEGIPIKRWLSRQNSLIGFPASTQDTTVALKALSEFA--ALMNTERTINQVTVGP 1212  
DB 1138 KDPAEKALPIIKWILISQNSNGGFSSTQDTYVIGLQALTKFAKYTGSGSGTMDIFFSSAGE 1197  
QY 1213 S-----SPSLPAAVQ-----PMANISANGGFPALCOLNVNRYKASGSSRRRSIQ 1259  
DB 1198 SKNTIKVNPENSLVLQTHDLPKSTRKVDFTKAGTGSAMVQSLTRYNLAEK-----E 1248  
QY 1260 NQAEFDLDVAVKENDKDILNVDLNVCTSF-----SGPGSGMALMEVNLISGFMVPS 1313  
DB 1249 KKPSPKATPYTKDIPNQLIYD--VCAEYVPLEADADKXSNMAMVEMELALPSGVGDS 1306  
QY 1314 ---ISLSETKVVEYDH--GKLNLYLDSVNETQF-CVNIPAVRNFKVSNTOVASIYDV 1367  
DB 1307 LGKIQADVRRKVRVETKNSDSSTVWVYFDSLTFPGDYRCLPLEASKAAVAKQCPASVLYDY 1366  
QY 1368 YEPRRQARVSNSEVKLSSCDLCSDV---QCCR 1397  
DB 1367 YDERKATEY--QVKSLLCDICGADCGEGCK 1397

RESULT 12  
ID 001717 PRELIMINARY; PRT; 1507 AA.  
AC 001717;  
DT 01-JUL-1997 (TREMBLrel. 04, Created)  
DT 01-JUL-1997 (TREMBLrel. 04, Last sequence update)  
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)  
DE Alpha-2-macroglobulin.  
OS Limulus sp.  
OC Eukaryota; Metazoa; Arthropoda; Chelicerata; Merostomata; Xiphosura;  
OC Limulidae; Limulus.  
OX NCBI\_TaxID=51645;  
RN  
RP SEQUENCE FROM N.A.  
RX MEDLINE=97175055; PubMed=9022715;  
RA Iwaki D., Kawabata S., Miura Y., Kato A., Armstrong P.B.,  
RA Quigley J.P., Nielsen K.L., Dolmer K., Sottrup-Jensen L., Iwanaga S.;  
RT "Molecular cloning of Limulus alpha-2-macroglobulin.",  
RL Eur. J. Biochem. 242:822-831(1996).  
DR EMBL; D83196; BAA19844.1; -.  
DR HSSB; P01024; IC3D.  
DR InterPro; IPR002890; A2M.N.  
DR InterPro; IPR002048; EF-hand.  
DR InterPro; IPR001599; MacroglloblnA2.

DR Pfam; PF00207; A2M; 1.  
 DR Pfam; PF01835; A2M; N; 1.  
 DR PROSITE; PS00477; ALPHA 2 MACROGLOBULIN; 1.  
 DR PROSITE; PS00018; EF HAND; 1.  
 DR SEQUENCE 1507 AA; 168490 MW; 88F4013495F63CED CRC64;  
 Query Match 22.3%; Score 1638; DB 5; Length 1507;  
 Best Local Similarity 28.9%; Pred. No. 7.4e-91;  
 Matches 449; Conservative 270; Mismatches 559; Indels 278; Gaps 43;  
 28 FLVATPAGIRRGQVNTIGVELLEHCPQSVYTKALLTASMLTYS-----VLEAG 78  
 29 FILTAPSLTPGKSLTILHLFD-----IKNGFLRGVCKQDDGVVAATEV 76  
 79 VFEK---GSFKTLTPS-----LPLNSADEIYELRVGTODEILFNSSTLSPEFTRIS 130  
 77 SFNDNPSSSIQILTPSGVEVKRPKLANSY-----SSPSNDPFPEKD--INHHKXLI 130  
 131 VFIOTDKALYKPKQEVKFRIVTLPSDFK--DYKTSMLILKDPKSNLILQMLSQOQSLGV 188  
 131 VFVQTDKPLKPKQGVKVRILPTTPDLKLVKETIGSFQIENPDGIVLGYPMLSPFAGI 190  
 189 ISKTFOLSHRILIDMSIQOVNDQTYQSFQSEVYLPRFEVTLQTPLYCSNMKSLING 248  
 191 AOFELALPDEPTYGMMWRKGNIEDTEIYENFEVKEVYLPKFEVKITPEPSYLLTNADSIW 250  
 249 TITAKYTGKPVKGDVTLTFLPLSFMGKKNIITKTPK-----INGANFSPNDE 297  
 251 KICQIYTGQVPEGTFAETVAVKKNKEGVPIHKEGLDGLDLYVASSA--LGNREQ 309  
 298 EMKQVMSNGLSEYLDLSSPGVEILTTVESVTGSRNVSTVFPKHDIIEFFDYT 357  
 310 RL-----SYRANMFALVEVTEKGTGIMKMATDRIYRTSPINIMYLEPT 352  
 358 T---VKPSLNFATVVKTRADGNQTLIEERRNNVITVQNTVTEWSSGNSQCKMEA 414  
 353 SGKYLKPLKPLKPLKPLKPLKPLKPLKPLKPLKPLKPLKPLKPLKPLKPLKPLKPLK 406  
 415 VQKINTVYPOSGTKIEF--PILEDSEIQLKAVFL-----GSKSMAVYSL--FKS 462  
 407 CKE--FTSDEAGIILKFTVPPOTPDITSPRFKAKLQYGGKDGKLNQPHSFTVSSWYS 464  
 463 PSKITIOLKTRDENIKVSGPEL--VVSNGRIKELSYNVVSGQLVAVGK----- 511  
 465 PSQSHLOLEPITEIECGKPLTVKFKYTTGEBKKQKFEYQIMANFIYDGSFHEPLIS 524  
 512 -----ONSTMFSLTPEN--SW-----TPKACVIV 533  
 525 EDKSGLTDETYLPIDVATSLNPNBEPENNVIIVPHIGETSLTLIPSEMMNSAKILV 584  
 534 YIISDGEIISDVLIKIPQVLFKKNKIKLYMSKVAEPSEKSLRISTVOPDSIYGIYAVD 593  
 585 FYVDEDETVADSTKITYKKLKRNVGLKRGEEKVLFGASTIQLTAS--PYSIGIAYD 643  
 594 KSVNLMNASNDITMENVHLEL-----ELVNTGY 622  
 644 KSVHILSSDRKITEEVEYKLGHDYWPQKATSDYKCYCEDYKQTEGHEGSSFGSGFT 703  
 623 LGHMANSPAVQEGGLVTLTANL-----TKDYIDGVYNAEAEAFMEENESHIV--- 673  
 704 STYNLDITAVDEGLVVISDMELETRPCKSGDEGPCPOQDVAFAAPQANRIGGG 763  
 674 -DIHDFSLG-----SSP--HYRKHPEPTWILDTNMGRIYQBEVAVPSSITSWATGF 725  
 764 GEAGGFGGGLRKTINKPVEIRITVFPETWLMELONIGATGELSKRIPIHTTEWGSAT 823  
 726 VISEDGLGLTTTPEVLQAPFFIPLNLPYVIRGEFALETITFNYLXADAEVKYIE 785  
 824 CISETGIGVSEA--ATYKGFQPFVSVFTLPYVIRGEKVPITVYVNYLSCELPKISLE 882  
 786 KSKDFDILMTSSSEINATGHQOTLLVPSBDGATVLPTRPHLGEIPITVYALS----- 838  
 883 QSDKF-----EMQNDTNSYTS CVCGKSGSDTTRMWMIKPRSLAGVNLTVGASIPNEAIC 935

QY 839 -----PTASDAVYTMILVKAEGIEKSYSSOSILLDTLNRLOSTLTKTISFSPPTVYG 891  
 DB 936 GNQDGYTVARDAATRQLLVEPEGPFKEDTWSTFACPKQONGFT--ATSDLLPDLVED 994  
 QY 892 SERVQITLIGDVLGSEINGLASLIRMPYCGGEONMINEFAPNIYILDTYTKKQTLTNLKE 951  
 DB 995 SARGVYISGDDLMGPAIKLIDHLVRLPTCGEONMVKFVFNIFVLDYLTATGSIYTSIKE 1054  
 QY 952 KALSPKQGYORELALQREDSGFSAGNVPDGSSTLSAFVLRCEFLADPYIDIDONVLA 1011  
 DB 1055 KALNNKRYGAKQNRHDPDGSISAFGNPKQGNLFLTFVYRSFQALERFILINXKLN 1114  
 QY 1012 RIYTMKQKQKSGNEFMDPGRVHSELQGC---NKSPTLITAVYTSLL--GYRKYQPN 1065  
 DB 1115 ETENWILNQRNSGCRKIKGLFNSALKGISNDETPAFLTAYVILISLEAGYKQ---- 1170  
 QY 1066 IDVQESIHFLESFSPSGIS-----DNYTLALITYALSSVSPAKKALNLTARAQ 1117  
 DB 1171 -----ETVYIDGISCLEALSNPSTYSLALFYAATSLAGHPAKQYLAKLEERAIT 1220  
 QY 1118 EGGMOFWVSESLDSWQPSRLDIEAAYA---LASHFLQFQTSBGIPIRMWLSQRNS 1174  
 DB 1221 EGGKTFWKSPPS--GRYWGNSIGVELAGYAVLTLQHGCAVLAQVTFIIRMLAQQNY 1278  
 QY 1175 LGGFASDTDTVVALKALSEFALMNTERTNIQVTVTPSPSPPLAVVQ----- 1222  
 DB 1279 RGGFSTQDTVIALQMSKFATIIYDELDELVGEVSSGFEKKIMLTCONSILMOTFRLO 1338  
 QY 1223 --PMAVINANGFGFALICQNLVYVYVYKASGSSRRRSIONQAFDLDVAVKEN--KODLN 1278  
 DB 1339 TVSPVDFEATGSGCLVQTSIRYVNTP-----PPRGFHELVYVYKGLYRDCIN 1389  
 QY 1279 HVDLNVCTSPSGPGR--SGMALMEVNLSGFMVPSAIS-----LSBTYKVEYDHQKLN 1332  
 DB 1390 -AHIACTVKYDGGVGNNAVLEMKVSGWIPDESIKIVIDREBLNRYEDVGNQML 1448  
 QY 1333 YLDVSNVETFCVNIIPAVRNFKVSNITQDASVSIYDYVEPRROAVRSYNSSEVLSKSD 1388  
 DB 1449 YPSELTDQMLCFMFLMEQDIEVQETKPATIRLYDYELQEVVTSYSID---ENCE 1501  
 RESULT 13  
 Q9UND7  
 ID 09UND7 PRELIMINARY; PRT; 1884 AA.  
 AC 09UND7;  
 DT 01-MAY-2000 (TREMELrel. 13, Created)  
 DT 01-MAY-2000 (TREMELrel. 13, Last sequence update)  
 DE 01-MAR-2003 (TREMELrel. 23, Last annotation update)  
 GN Hypothetical protein KIAA1283 (Fragment).  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Brain;  
 RX MEDLINE=20039619; PubMed=10574462;  
 RA Nagae T., Ishikawa K., Kikuno R., Hirosewa M., Nomura N., Ohara O.,  
 RT "Prediction of the coding sequences of unidentified human genes. XV.  
 RT The complete sequences of 100 new cDNA clones from brain which code  
 RT for large proteins in vitro.",  
 RL DNA Res. 6:337-345 (1999).  
 DR EMBL; AB033109; BAA86597.1; -.  
 DR HSSP; P01024; IC3D.  
 DR InterPro; IPR002890; A2M\_N.  
 DR InterPro; IPR002350; kazal.  
 DR InterPro; IPR001599; Macrogloblina2.  
 DR Pfam; PF00207; A2M; 1.  
 DR Pfam; PF01835; A2M; N; 1.  
 DR Pfam; PF00050; kazal; 1.  
 DR SMART; SM00280; KAZAL; 1.

DR PROSITE: PS00477; ALPHA\_2\_MACROGLOBULIN; 1.  
 KW Hypothetical protein.  
 FT NON TER  
 SQ SEQUENCE 1884 AA; 206524 MW; D3C078F26B951D1A CRC64;  
 Query Match 21.0%; Score 1542.5; DB 4; Length 1884;  
 Best Local Similarity 27.4%; Pred. No. 7.1e-85;  
 Matches 494; Conservative 257; Mismatches 555; Indels 497; Gaps 57;

5 PLTAAHLICVC-----TALAVALAPGRPLVTAGIIRPGSNVITIGELLEHCQSYTYA 60  
 7 PLPLILLILLASADGVRAPAPAG--YLLAASVFRAGVEEYISVITIP-NSREYTYA 63  
 61 ELKTAASNLVSVLEAG--VEFKSGFKTLLPLSPINSADIEYELRVTGR---TODEILF 116  
 64 QLVAGE-----PVVSGAAILDKGTIK-LKPTGLAGQA-----LKVWRGMQAEGEPLF 114  
 117 SNSTRISFETKISVITIQDKALKYKQVYKRIYTLFSDPFKYKTSNLILKDKPSNLI 176  
 115 HNOTSVTVDRGASVFIQTDKPYRRPQHVLISIFVSPMLRPNNEKLEAYIILDPGSHM 174  
 177 QQWLS--QSGDLGIVSTFQLSHPLIGDMSIQVQVNDQYVQSGVQSEVVLPRFEYTLQT 235  
 175 IEMRHLPCCCGITNNSFPLSDQVLGEWTFEVMQGHAYNNSFEVQKYLPRFELLIDP 234  
 236 PLVCSMNSKHLNLTITAKTYGKPVKQDVT--LTFLPLSFGKCK--KNITKTFKINGSAN 291  
 235 PRITQDLACEITGAVARTFGKPVAGALTINNTVAGVGVSHVEGRPLRTKILIGSD 294  
 292 FSEFNDEEMKNVWDSNGLSBYLDLSP---GPVELITVTVESVTGISRNSTNVFFKOH 347  
 295 FDLGVADM-----IPADVPEHGRGVSIWAMVT--SVDSQO----- 329  
 348 DYIIIEFDYTV-----LKPDLNFATVATVRADNOCLTEERRNNVYI 391  
 330 ---VAADDSTPVQRQLVDIRYSKOTRKQFPGIAYGVKELSLYDPSD-----AEGTVV 380  
 392 TVTORVYTEWMSGNSGNQMEAVQKLN--YT---VPQSGTFKLEFPLEDSSE---LOLK 444  
 381 QI-----KAEITPNDNITYTSEVSGRLGVGEITISITPSAOHWLETK 423  
 445 AAVFLGSK-----SSMAVHSLFKSPSKTYIQLKTRDENIKVSGPFLVSGNKKLK-E 495  
 424 VMALNKGVPVGAQYLPYLSISWY--SPSQCYLQLQPSHPLOQGEAYSVSTCCNFT 482  
 496 LSTMVISRGQVAVGQ-----N 513  
 483 LYEVAVARGNIVLSGOOPAHHTTOQRSKRAPALEKEPIRLTHLSETPPAPAEADVCVT 542  
 514 STMFSLTPENSWTPKACVIVYIETDEGEIISDYLKI PVQLVFNKKIKLWMSKYKAPSRK 573  
 543 SLHLAVTP--SNVPLRLVLFVYRENGEVAADSLQAVETFEFNQSVITYSANETOPGEV 600  
 574 VSLRISVTOPDSIVGIVAVDKSVNLNANSDITMENVAHLELYNTG----- 620  
 601 VDLRIIAAR-GSCVCAAVDKSVYLLRSGPRLTPAVQFQELDEYDVSDSFGVSRDEGPF 659  
 621 -----YILGFEMNSPAVQECGLAVLTQ-ANITKQYIDGVYDINAETAEAR 663  
 660 MAGLTAQRRRRSVFPWPGITKDSGAFLETGLVMTDRVSLNHRQDGLY--TDEAVP 717  
 664 FMEENGHIVDIDFSLGSSPH-----VRKHPEFTWIMLDTNMGYRIYQ--EFEYTVPD 715  
 718 AQQPHIGSLV-----AAVPSRHPPTREKAKRTFFETWIMHCLNIDPSGEGSLVAVVD 772  
 716 SITSWATGVISEDGLGLTTPVELQAFQPFIFLNPYSVIRGEFALETITFNPK 775  
 773 SITSGEVAVALSTQGLGI-AEPLSLKTFKPFVDFMLPALIRIEGVKIPLSVYVNWG 831  
 776 DATEV--KYIEESDKFDILMTSSEINAGH-----QOTLVPSDGAIVLFPPIPTLU 827  
 832 TCAEVYMKLSVPGKIGF-----VGHPEKRVHTYKMKVAPGEAEPIMVWVLSFSDL 880

QY 828 GEIPITVTLAS-----PTASDAVYQOMILVKAEGIEKSYQ 862  
 DB 881 GLNNITAKALAYGDTNCCRODRSSKHPEENHADRVPVIGVDHVRSSVWVEAEGVPRAVY 940  
 QY 863 SIL-----LDLT-----DNRL----- 873  
 DB 941 SAFFCPSRVIHSPENKYEFOYVQRPRLTTFEDVAVRAHNDARVALSSGPODTAGMIEIV 1000  
 QY 874 -----QSTLKTLSF----- 882  
 DB 1001 LGHONTSMWISTSMKGPVASANTAKILSWDEFRTWISWRGGLIQVGHGPEPNSSVI 1060  
 QY 883 -----SFPNPTVGSERVOIT 898  
 DB 1061 VAWTLPRPEVQFIFGTSNGWGMGEFRIMRKMEVDESSEAFTLGVPHGALPGSSRRATAS 1120  
 QY 899 AIGDVLGSLINGLASLIMPYGCGEOMNINFAENIYIIDYLTKKKQLTQNTKEKALSFWR 958  
 DB 1121 IIGDWMGPETLNHLNLTLPFGCCGEOMNIHFAPNVFLKYLOKQOLSPEVERETDYL 1180  
 QY 959 OGORELLYOREDSFSAFGNYDPSGSTWLSAFYLRCGLEADPYIIDQNVLIHRTYMLK 1018  
 DB 1181 QGYRQLTYKQODGSYSAGERDASGSMWLTAFVKSFAQARSFLVDPRELAANKSVIT 1240  
 QY 1019 GHQKNGEFWDPGRVHSELQGNKSPLYTLAYITVSL--GYRKQYQPIDVESIHFL 1076  
 DB 1241 QQQQADSGFLAVGRVANKDIOGGIHGIVPLRAYVVALLETGTASEERGSSTDARHLE 1300  
 QY 1077 SEFSRGISDNTYALITYALSSVSGPRAKEALNMTWRBADEGGQMFVSSSKLSDSQ 1136  
 DB 1301 SAAPLAM-DPYSCALTYALTILRSAPAPEALRKLSLAIMRDVTHW-----SLSNSMD 1354  
 QY 1137 PR-----SLDIEVAVALLSHPLQFQFSEGIPIRMWLSROBNSLIGFAS 1180  
 DB 1355 VDKGTFLSFSDRVSQSVASAEVEMTAYALTYTLGDVAALPVVKWLSQQRNALGGSS 1414  
 QY 1181 TQDTTVALKALSEFAL-----NANTERTNIQVTVG--PSGPS 1216  
 DB 1415 TQDTVALQALAEVAILSYAGINITYSLASTNDYQTFELHRTNQVLOTAAIPSLPT 1474  
 QY 1217 PLAVVQMAVNIISANGFPAICQLVVYVVK---ASGSSRRRSRISQNGE----- 1262  
 DB 1475 GLFV-----SAKGDCCLMQIDVTYVNDPVAKPAFOLVLSQBEAQRPPPMAS 1526  
 QY 1263 -----AFPLDVAVKENKODLNHYDLNCTCSBGSBGRSGMALMEVNLISGFMPWS 1311  
 DB 1527 AABSGRGWPPADDDPPADQHOEYK--VMLEVCTRMHLAGSSNVAVLEVLGFRADI 1585  
 QY 1312 EALISLET-----VKKEVEDHGLNLYLDSV--NETQFCVNIIPAVRFFKVSNTQDASVS 1363  
 DB 1586 E--SLBQLLDKMKMKRYEAVAGRVLFFPEIIRSCITCVRFRLRCEVVGRTSALPVS 1643  
 QY 1364 IVDYEPRAQVRSYNSSEV-----KLSGCDLGSQDVQ-----GCRCEDEGASGSHHS 1410  
 DB 1644 VDYDEPAFEATRFYNVSTHSLARELCAGPACNCEVERAPARGGWFGESEGPAVAPBEG 1703  
 QY 1411 SVI 1413  
 DB 1704 AAI 1706

RESULT 14  
 DB 0812J3 PRELIMINARY; PRT; 1885 AA.  
 AC 0812J3;  
 DT 01-MAR-2003 (Tremblrel. 23, Created)  
 DT 01-MAR-2003 (Tremblrel. 23, Last sequence update)  
 DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)  
 DE Alpha-2 macroglobulin family protein VIP.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 NX NCBI\_TaxId=9606;

[1]  
 RN SEQUENCE FROM N.A.  
 RA Li Z., Wu X., Engvall E.  
 RT "Cloning and expression of VIP, a novel alpha 2 macroglobulin family  
 member".  
 RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AY101765; AA050084.1; -  
 SQ SEQUENCE 1885 AA; 206655 MW; 3FB48551FBF326D CRC64;  
 Query Match 21.0%; Score 1542.5; DB 4; Length 1885;  
 Best Local Similarity 27.4%; Pred. No. 7.1e-85;  
 Matches 494; Conservative 257; Mismatches 555; Indels 497; Gaps 57;  
 QY 5 PLTAHLAHLVCV---TAALAVAGPRLVAPGIIIRPGANVITGVLELHPSQVTVKA 60  
 DB 8 PLPLPLLLLSARQVRAAQAQPAQ---YLIAAPSVFRAQGEVIVSVITF-NSPREVTVQA 64  
 QY 61 ELKTAASNLVSVLEABG-VREKGSFKTLPLPLPLASADEIYELRVTGR--TODEILF 116  
 DB 65 QLVAGQGE---PVVQSGAILDKGTIK-LKVPITLARGQA---LLKTVGRGMQAEGPLF 115  
 QY 117 SNSTRSPETKRISVFIQTDKALKYKPKQEVKFRIVTLESDEPKYKTSNLIKDKPSNLI 176  
 DB 116 HNOISVTVDGASVFIQTDKPVTRPQRRVLISIFVSPNLRPVNEKLEAYILLDPGSRM 175  
 QY 177 QQMIS-QQSDLVGSKTFOLSHRPLGDMISQOVNDQTYXSFQVSEVYLPRKEVTLQT 235  
 DB 176 IEMHLKPFCCGINTMSFPLSDQVLEGMFIFVEMQGHAKYKSFVQKTVLPKRELLIDP 235  
 QY 236 PLYSKMSKHLNGITTAQYVYGVKPKVDVT--LFLPLSPFGKK--KNITKTFKINGSAN 291  
 DB 236 PRYIODLACETGTVRATYFGKRPVAGALITNTMNGVGVSHVGRPVLTETKILASQD 295  
 QY 292 PSFDEEMKVMDSNGISEYLDLSP---GPVEILTVTESYTGISRNSTVVFVKQH 347  
 DB 296 FDCIVRDM-----IPADVPEHFRGVSIAMAVT-SVDGQQ----- 330  
 QY 348 DYIIEFFDYTV-----LKPSLFTATVKTTRADGNQLTLEERNVVI 391  
 DB 331 ---VAFDSTFVQAGQVLDIRYSKOTRKQFKGLAVGVKVELSYDGSB---AEGTV 381  
 QY 392 TWTORNTYEWGSGNSGNQKMEAVOKIN-YT--VPQSGTFKIEPILLESSE---LQK 444  
 DB 382 QI-----KAELTPKONITSEVVSQGLVGFEPISPTSAQHWLETK 424  
 QY 445 AYFISGR-----SSMAVHSLFKSPKTYIOUKTDENIKYKSPRELYVSGNKRLK-E 495  
 DB 425 VMAINGKPVGAQVLPSTYLSGSMY-SPQCYLQLOPSPHPLQVEEAVFSYKSTPCNFT 483  
 QY 496 LSYVWVSRGOLVAVGKQ-----N 513  
 DB 484 LYVEAARAGNIVLSGQCPRAHTTQORSKRAALKEKPIRLTHLSTTEPPRAEAVDYCVT 543  
 QY 514 STMSSTLPEPNSWTERKACVIVYIEDDEIISDVLIKIPQVLFKNKIKLYMSKYAEPSEK 573  
 DB 544 SLHLAVTP--SMVPLGRLLVFEYRENGEGVADSLQFAVETFEFQVSVTASANETQGEV 601  
 QY 574 VSLFISTVQPSIYGIYAVDKSVNLMAASNDITMENVHELELYNTG----- 620  
 DB 602 VDLIRIRAR-GSCVCAVADVSVYLLRSGFRLTPAQVFOELEDDVSDSEGVSRDEDPFW 660  
 QY 621 -----YILGMENNSFAVFOECGLMLTD--ANTLKDYIDGYVDAEYAE 663  
 DB 661 MAGLTAQRRRRSSVFPWPMWGITKDSGFAFTETGLVMTDRVSLNHRDDGLY--TDAV 718  
 QY 719 AFQHTGSLV-----AAVPSRHPRTKRTKRTPEPTMIMCNISPSGSGTLSTVAVPD 773  
 QY 716 SITSWATGFIYSEDGLGLTTTVEVLAQFQPFIFLNLPSVYRGEBFALEITITNYLK 775  
 DB 774 SITSMGVAVALSTSGIGI-AEPSLLKTFKFPFVDFMLPALIIRGEQVKIPLSVYVYMG 832

QY 776 DATEV--KVIEKSKDFILMTSSEINATGH-----QOTLVPSDEGATVLPFIRPTHL 827  
 DB 833 TCAEVYWKLSVPRKGIQF-----VGHPEKRYHTKKMCVAPGEAEPIVWLSFSDL 881  
 QY 828 GEIPIYVTLAS-----PTASDANTQMLIVKAEIGEKSYSQ 862  
 DB 882 GLNNITAKALAYDNTCCRDGRSSKHPENHADRAVPIGVDRRSVWEAEVPRAYTY 941  
 QY 863 SIL-----LDLT-----DNL----- 873  
 DB 942 SAFFCSERVHISTPKYEFQYQVRLRFLRPVAVRAHNDARVALSSGQDTAGMIEIV 1001  
 QY 874 -----OSTLKLTSF----- 882  
 DB 1002 LGGHQTRSWISTSKMGEVPVASHAKILSMDFRTFMIWMRGGLIYGVHGPEPSNESVI 1061  
 QY 883 -----SPPNPTVTSERVQIT 898  
 DB 1062 VANTLPPPEVQIFGISTGMSKGEFRIMRKEVDESSEAFPLVGPBGALPGESEATAS 1121  
 QY 899 AIGDVLGPSINGLASLIRMPYGCQEQNMINFANIVYILDYLTKKOLTDNLKELKALFMR 958  
 DB 1122 IIGDVGPTLNLNLLRPLFGCGEQMILHAPNVFLKLTQKTQGLSPEVERETDYL 1181  
 QY 959 QGYQRELLYQREDGSFAFGNVDPSGSTWLSAFVLCFLEADPYIDIDONVLRHTYTWLK 1018  
 DB 1182 QGYQRLQYTRQOGGSYAFGERDASGSMWLTAVLKSFAQARSFIFVDPRELAASKWIT 1241  
 QY 1019 GHQKSGEFMDPRVHSELQGNKSPVLTATVYLSL--GYRKQPNIDVOESIHFLF 1076  
 DB 1242 QQQQADGSFLAVGRVANKIDGGIHGIVPLTAVVVALLETGASSEERSTDKAHFLE 1301  
 QY 1077 SEFSRGISDNYTALTITVYALSVSGPAXKALMMLTMRABQEGQMFVWSSEKLSDSMQ 1136  
 DB 1302 SAAPLPM-DPYSCLATYTLTLRSRAAPBALKLSLAIMRGVTHM-----SLNSMD 1355  
 QY 1137 PR-----SIDIEVAAYALLSHFLOFQTSBGLPIMKLSRQNSLGGPAS 1180  
 DB 1356 VDKGTFLSFSDRVQSQVSAEVEMTAVALLTYLLDVAALAPRVVWKLSSQRRNALGFS 1415  
 QY 1181 TQDTYALKALSFAL--MTRRTNIOQTVTC--PSSPS 1216  
 DB 1416 TQDTYALQALAYALLSYAGINTVSLASTLWDYOETFEHLRTQKVLQTAIPLSLPT 1475  
 QY 1217 PLAVQPMVNIISANFGFALCOLANVYVYK--ASGSSRRRSIONOE----- 1262  
 DB 1476 GLFV-----SAKEDCCIMQIDVTYVPPDVAPARQQLVSLQEPHOGPPEMPAS 1527  
 QY 1263 -----AFDLVAVAKENDLNHYDLNCTSPSGRSGMALMEVNLISGFMVPS 1311  
 DB 1528 AAGSGRDMPRADDDPADQHHQYK-VMLEVCTRMHLAGSSNMAVLEVPILSGFRADI 1586  
 QY 1312 EALSISLET-----YKQVEYDHGKANTLYDS-NEQPCVNIIPAVNFRVSTQDASVS 1363  
 DB 1587 E--SLQQLDKRMKMKRYEAVAGRVLFFYFDELPSSCLTCVRRALRECVGRTSALPVS 1644  
 QY 1364 IVDYEPRAQVASYNSEV-----KLSSCDLCSQVQ-----GCRPCDASGSHHS 1410  
 DB 1645 VDYIEPAFAITFVNVSTHSPILARELCAGPACNEVERAPARQPGMFPBESGPAVAPBEG 1704  
 QY 1411 SVI 1413  
 DB 1705 AAI 1707

RESULT 15  
 ID 060486 PRELIMINARY; PRT; 1476 AA.  
 AC 060486;  
 DT 01-NOV-1996 (T-EMBLrel. 01, Created)  
 DT 01-NOV-1996 (T-EMBLrel. 01, last sequence update)  
 DT 01-MAR-2003 (T-EMBLrel. 23, last annotation update)  
 DE Alpha-macroglobulin precursor.



OS Cavia porcellus (Guinea pig).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Hystriognathi; Cavidae; Cavia.  
 OX NCBI\_TaxID=10141;  
 RN  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=97164019; PubMed=9010766;  
 RA Iwasaki H., Suzuki Y., Sinoheara H.;  
 RT "Cloning and sequencing of cDNAs encoding plasma alpha-macroglobulin  
 and murinoglobulin from guinea pig: Implications for molecular  
 evolution of alpha-macroglobulin family.";  
 RT J. Biochem. 120:1167-1175 (1996).  
 RL EMBL; D84336; BAA12316.1; --  
 DR HSSP; P01023; 1BVB.  
 DR InterPro; IPR002890; A2M\_N.  
 DR InterPro; IPR001599; MacrogloblnA2.  
 DR InterPro; IPR000531; Tonb\_boxc.  
 DR Pfam; PF00207; A2M\_1.  
 DR Pfam; PF01835; A2M\_N; 1.  
 DR PROSITE; PS00477; ALPHA\_2\_MACROGLOBULIN; 1.  
 DR PROSITE; PS00430; Tonb\_DEPENDENT\_REC\_1; 1.  
 DR Signal.  
 KM SIGNAL.  
 FT CHAIN 1 23 POTENTIAL.  
 FT SEQUENCE 1476 AA; 163614 MW; 0550693A068DC5C CRC64;  
 SQ  
 Query Match 20.3%; Score 1491; DB 11; Length 1476;  
 Best local Similarity 28.8%; Pred. No. 6.7e-82;  
 Matches 451; Conservative 292; Mismatches 524; Indels 300; Gaps 60;

4 PRLTAHLVCCTAALAVAGPRPLVTAPGIRPGANVTIGELLEHGPSQVTVARL 62  
 11 PVL-----LLILAADASISGKPGQVWVIVSLHSGPEKIC-LTQINETVTVKASLD 65  
 63 -LKTASNLVSVLEAGVFEKSGFKTLTLPSEPLNSADEIYELRVTKGTODELFESNST 121  
 66 TIRENGSLPFMNMAEKLQCVAF---TVQSPYPAVWELTVVEVGPTHG--FRSRKT 119  
 122 LSFETRISVFIOTDKALYKPKQEVKFRITVLPSPDKPKYKTSLN-----ILIKDPKSNLI 176  
 120 VLAKSDSLVFVOTDPRYKPGQTVAKAVVSLDENFRP-----LNEIPLFIQDPFGANV 175  
 177 QQWLQQOQSDLGIVISKTFOLSHPILDWSIQV-QVNDQTYQSFQVSEVYLPEFEVTLQT 235  
 176 MQQNLKLEKGLQLSFLSSSEPLDGSYVVAHKGSGRMHSHFTVEFEVLPPEFEVQVM 235  
 236 PLYCSMNSKHLNCTIAKTYGKPVGVDTLTL-----PLSFWKKKNIYTFE-KING 288  
 236 PKKTIIEQEFVTVSGRYTYGKPVGNTIMSICRYNNPNSACLSEB--SRAFCCKVQ 292  
 289 SAN-----FSFNDEMK-----NVMDSNGLSEYLDLSSPGVEILLTVYT 328  
 293 QANSGCFFQOVKTNFOLRREKYEMLRVEAKIREEGTG---VOLGTGSEITATYT 348  
 329 ESVTGISRVNSTVNFKQHDYILFFDYTTVLKPSLNFTATVAVKTAADGNOLLEERRNN 388  
 349 K-----LSFVKADSVYRAGVPFGQVRL--VDKKNVMPHKKM-- 383  
 389 VATTQRYNTYEWSSNGNQMEAVQ-KINTYPOSGTFKIEPILDESS---ELQL 443  
 384 ITITAEAVY-----HSNATTDENGLVQFSINTNMIGTSLNQVHGKSTNCYDQWML 438  
 444 KAYFLGSKSMVAHSLFKSPSKYIOLKT-----RDENIKVSGPFELVUSGNRLKE 495  
 439 EANEASHHTANAFAFSI-----SRSFVLELPQLGLPCHQOTQFPA---HYILKG-OLKE 489  
 496 L--SYNVSRGQLVAVG-----KONST-----MFSLTPEMSWTPKACVIVYIEDGETI 543  
 490 LVFYYVYIMAKGIVQSGTVYLVSEQNTKCHPSVSPVESDLARVAVLIYALIPSEI 549  
 544 SDVLKIPVOLVFNKIKLYMSKYKABPSEKVSRLISVTPDSTVIGIVAVDKSVNL----- 598  
 550 AUSAQYVNECLDNKYNLSFSEGSILPASKTHLRVTAS--POSTICALRAVDQSVTLRKPA 608

599 -MNAEN-----DIT-----MENVHELELYNTGYLGMF-----MNGFA 631  
 609 VLSASSYVALLPVDOLTPGPRGLGQEBNDEBCVSLYNT-YIDGLYPPERINENKMYG 667  
 632 VFQECGLVLTIDANLTQYIDGVNDAEYAEERFMEENEGHIVDHFSLGSSPH----- 685  
 668 FLKDKLAKFNTKLOKPOLCAHVQKFEVPTMAVSYSSSS-----SFRSGPRVPVAVG 720  
 686 -----VRGHFRETW---LDTNMGYRIQFEPVTVPSISWATGVISPD 730  
 721 IATYSEPPKETVRYSPETIWDKVTDSG--VAEVEVTPDTITTEWKAACLSND 777  
 731 LGIGLTTTPVELOAFQPEFIPLNLPYSVIRGEEFALETIIFENKDATEVKYIEKSPK 790  
 778 TOLGISPT-ASIRAQPFVELTMYSVIRGEAFILKATVANYLPDCLRISVHLEASRKF 836  
 791 DILMTSSBINATGHQOITLVSESDATVLPFIRPTHLGEIPTVA----- 836  
 837 LAEPRAKE-----QESYCVCGNERQVSWVTPKSLGNVFTVSAEALSESSELCNKT 890  
 837 LSPT--ASDAVYQMLVVAEGIEKYSQSILDLTDNLOSTLKTLSFPNVTGSE 894  
 891 VPTTGKQDTIILKPLVEPEGIEKETWLSIRSDTVS--EKHLDELPSNVYQDSAR 947  
 895 VQITAGVLPSPINGLASLRMPYCGEONMINAPNITYIYDYLTKKQOLTNLKERAL 954  
 948 ATVSLIGDILGSAMONITLQMPYCGEONNVLPAPIYVLDYLNFTQOLTPDKSKAT 1007  
 955 SFMRQYQRELLYQREDSFSAFG-NV-DPSGSTLSAFVLRCELEADPYIDIDQNVLR 1012  
 1008 SYLSGYQRLNYSKRDGSGTFGENYRGQGNMTLTAFLVKTFSQARKYFIDAHITQ 1067  
 1013 TYTLWKHQKNSGEWDEGRVYHSELQGNKSPVTLTYIYTSLSIGYRKQPNIDVQSI 1072  
 1068 ALSWLSOKQKNGCPSWSSGSLNMAIKGVDEISLSAYITITALEMSLPDHPVVRNAL 1127  
 1073 HPLES-----EFSRGISDNTYLTALITYALSVSQ-PAKEXALNMLTIRABEGGMQFV 1125  
 1128 FLESAMKSAKEGTG-SHYITKALAVAFALAGQERKKEILMSLEBGVDENSLHMA 1186  
 1126 SSES-KLSDSW-----QPSRLDIEVAAYALLSHFLOFQTSSEGP-----IMRWLSR 1170  
 1187 RPQKRVSBGFLFKSQARSABEEMTSYVLAVL-----TARAPFPEDLTSATDYNWTK 1242  
 1171 QNSIGFASLTODTVALKALSEFAA--LANTERTNIOVTVGSSSP-----LAV 1220  
 1243 QONSHGYSSTODTVALHALSKYAAATFRTTEKA-AQVITKSGTFTNEFVNHNRL 1301  
 1221 VQPMV-----NISANGPRAICQMLVVNV-----KASSGSRRR 1256  
 1302 LQOVSPLPVSDSYITTVTBGNGVYIQTSLKNVSEKCTPPRALAEFTVPOACDGPKN 1361  
 1257 STONEAEPDLVAVKENKDINHVDLVNCTSPSGRSGMALMEVNLISGF--MVP-BA 1313  
 1362 SFQ-----ISLNVSYGSRP-VSNMAIVDKYMSGFIPLKPIYKN 1400  
 1314 ISLSETVKVEYDHGKMLYLDVSNETOFCVNIPAVNFKVSNODASVIVDYIEPRQ 1373  
 1401 LEKSEHISRTESNNHVLIDKVSNQTLISLFFVQDIEVRDLKPAIKKYDYETNEF 1460  
 1374 AVRSYNS 1360  
 1461 ALAEYHA 1467

Search completed: January 15, 2004, 18:14:03  
 Job time : 64 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: January 17, 2004, 00:12:55 ; Search time 5629 Seconds  
(without alignment)  
6165,712 Million cell updates/sec

Title: US-10-020-095-4  
Perfect score: 7348  
Sequence: 1 MGGPPLTAHLICVCTAL.....HSSVIFCFKLYFMELML 1428

Scoring table: BLOSUM62  
Xgapop 10.0, Xgapext 0.5  
Ygapop 10.0, Ygapext 0.5  
Fgapop 6.0, Fgapext 7.0  
Delop 6.0, Delext 7.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEL=frame+ p2n.model -DEV=xlp  
-O=/cgn2.1/USPFC pool.P/US10020095/runat.15012004.163822.18713/app\_query.fasta.1.1607  
-DB=BST -QFMT=fastap -SUFFIX=rcs -MINMATCH=0.1 -LOOPEXT=0 -LOOPEXT1=0  
-UNITS=bits -START=1 -END=1 -MATRIX=biosum62 -TRANS=human40.cdi -LIST=45  
-DOCALLIGN=200 -THR SCORE=PCT -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=pco -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000  
-USER=US10020095 @CGN 1.1 4337 @runat.15012004.163822.18713 -NCPU=6 -ICPU=3  
-NO MMAP -LARGEBUFFER -NEG SCORES=0 -WAIT -DSPPLOCK=100 -LONGLOC  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :  
1: em\_estba:\*  
2: em\_esthum:\*  
3: em\_estlin:\*  
4: em\_estmu:\*  
5: em\_estrov:\*  
6: em\_estrpl:\*  
7: em\_estro:\*  
8: em\_hlc:\*  
9: gb\_est1:\*  
10: gb\_est2:\*  
11: gb\_hlc:\*  
12: gb\_est3:\*  
13: gb\_est4:\*  
14: gb\_est5:\*  
15: em\_estfun:\*  
16: em\_estcom:\*  
17: em\_gss\_hum:\*  
18: em\_gss\_inv:\*  
19: em\_gss\_pln:\*  
20: em\_gss\_vrt:\*  
21: em\_gss\_fun:\*  
22: em\_gss\_mam:\*  
23: em\_gss\_mus:\*  
24: em\_gss\_pro:\*  
25: em\_gss\_rtd:\*  
26: em\_gss\_phg:\*  
27: em\_gss\_vrt1:\*  
28: gb\_gss1:\*

29: gb\_gss2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	5343	72.7	4485	11 AK029247	AK029247 Mus muscu
2	5304	72.2	4725	11 AK036799	AK036799 Mus muscu
3	1480	20.1	4641	11 BC044072	BC044072 Xenopus 1
4	1461	19.9	941	13 BX391819	BX391819 BX391819
5	1406	19.1	1201	9 AL554092	AL554092 AL554092
6	1405.5	19.1	4667	11 BC049868	BC049868 Mus muscu
7	1280	17.4	879	13 BU517328	BU517328 Mus muscu
8	1117	15.2	667	12 BQ028040	BQ028040 UI-H-CO0-
9	1022	13.9	1201	9 AL576908	AL576908 AL576908
10	1007	13.7	1495	11 AK042169	AK042169 Mus muscu
11	1006	13.7	978	13 BX386457	BX386457 BX386457
12	979.5	13.3	836	10 EG173799	EG173799 602333853
13	972.5	13.2	638	13 BU624356	BU624356 UI-H-RG1-
14	941	12.8	648	10 BB613975	BB613975 BB613975
15	912	12.4	634	9 AA418644	AA418644 zvc3f07.r
16	902.5	12.3	613	14 CB158337	CB158337 K-BST0217
17	895	12.2	919	13 BX370169	BX370169 BX370169
18	870	11.8	988	13 BX386456	BX386456 BX386456
19	849.5	11.6	586	13 AA284110	AA284110 zc36a07.T
20	833	11.3	473	13 BX279537	BX279537 BX279537
21	831	11.3	537	14 CA417651	CA417651 UI-H-FE0-
22	830	11.3	515	13 BX472602	BX472602 DKFZP686C
23	829	11.3	544	10 BE143588	BE143588 MRO-HTR016
24	759	10.3	853	10 BE573834	BE573834 601331813
25	750.5	10.2	835	14 CD242054	CD242054 AGENCOURT
26	746	10.2	490	12 BI975931	BI975931 484508 MA
27	738	10.0	437	9 AA132252	AA132252 z006h07.r
28	732	10.0	456	9 AI278292	AI278292 q178d07.x
29	731	9.9	480	12 BM537249	BM537249 ha80a02.g
30	692	9.4	684	9 AV245523	AV245523 AV245523
31	687	9.3	806	12 BG919720	BG919720 602822407
32	671	9.1	722	14 CA420610	CA420610 UI-H-FG0-
33	661	9.0	511	14 CA893496	CA893496 BU179D10-
34	660	9.0	620	13 BQ265550	BQ265550 NISC ffl3
35	649	8.8	512	14 CA566520	CA566520 K0402C12
36	645	8.8	395	12 BG983552	BG983552 IL5-CN006
37	636	8.7	717	13 BQ775821	BQ775821 UI-H-FH0-
38	635	8.6	635	14 CA429478	CA429478 UI-H-FH1-
39	633.5	8.6	673	10 BB639005	BB639005 BB639005
40	626	8.5	476	13 BQ551040	BQ551040 H4067H10-
41	614	8.4	687	13 BQ773484	BQ773484 UI-H-FE0-
42	612.5	8.3	432	14 CB158287	CB158287 K-EST0217
43	610.5	8.3	819	13 BU434307	BU434307 603836006
44	609	8.3	431	10 BE006130	BE006130 RC4-BW012
45	597	8.1	652	12 BQ020033	BQ020033 UI-H-ED0-

#### ALIGNMENTS

RESULT 1  
LOCUS AK029247  
DEFINITION Mus musculus 0 day neonate head cDNA, RIKEN full-length enriched library, clone:4831440K17 product:weakly similar to alpha-2-macroglobulin (fragments) [limulus polyphemus], full insert sequence.  
ACCESSION AK029247  
VERSION AK029247.1 GI:26081244  
KEYWORDS HTG; CAP trapper.  
SOURCE Mus musculus (house mouse)  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

REFERENCE	1	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
AUTHORS		Carninci, P. and Hayashizaki, Y.
TITLE		High-efficiency full-length cDNA cloning
JOURNAL		Meth. Enzymol. 303, 19-44 (1999)
MEDLINE		99279253
PUBMED		10349636
REFERENCE	2	Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
AUTHORS		Itch, M., Kono, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
TITLE		Normalization and subtraction of cap-trapper-selected cDNAs to
JOURNAL		prepare full-length cDNA libraries for rapid discovery of new genes
MEDLINE		Genome Res. 10 (10), 1617-1630 (2000)
PUBMED		20499374
REFERENCE	3	Shibata, K., Itch, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,
AUTHORS		Kono, H., Akiyama, J., Nishi, K., Kitsuma, T., Teshiro, H., Itch, M.,
		Sun, N., Ishii, Y., Nakamura, S., Hazama, T., Nishino, T., Harada, A.,
		Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,
		Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watanabe, K.,
		Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsunaga, S., Kawai, J.,
		Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
TITLE		RIKEN integrated sequence analysis (RISA) system--384-format
JOURNAL		sequencing pipeline with 384 multicapillary sequencer
MEDLINE		Genome Res. 10 (11), 1757-1771 (2000)
PUBMED		20530913
REFERENCE	4	Kawai, J., Shingawa, A., Shibata, K., Yoshino, M., Itch, M., Ishii, Y.,
AUTHORS		Arakawa, T., Hara, A., Fukunishi, Y., Kono, H., Adachi, U., Fukuda, S.,
		Altawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamataka, I.,
		Salto, T., Okazaki, Y., Gotojori, T., Bono, H., Kasukawa, T., Salto, R.,
		Kadota, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T.,
		Fischmann, W., Gaasterland, T., Gissi, C., King, B., Kochiwa, H.,
		Kuehl, P., Lewis, S., Matsuo, Y., Nakado, I., Peeble, G.,
		Quackenbush, J., Schriml, L. M., Staudl, P., Suzuki, R., Tomita, M.,
		Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H.,
		Baldarini, R., Barish, G., Blake, J., Boffelli, D., Bojunga, N.,
		Carninci, P., de Bonaldo, M. P., Brownstein, M. J., Bult, C.,
		Flitcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D.,
		Hofmann, M., Hume, D. A., Kamya, M., Lee, N. H., Lyons, P.,
		McDonald, L., Nishida, J., Mazarrelli, U., Mombere, P., Nordone, P.,
		Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H.,
		Sato, K., Schonbach, C., Seya, T., Shibata, Y., Storch, K. F., Suzuki, H.,
		Toyo-oka, K., Wang, K. H., Wetz, C., Whitaker, C., Wilming, L.,
		Wynshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawaji, H., Kohlsaki, S.
		and Hayashizaki, Y.
TITLE		Functional annotation of a full-length mouse cDNA collection
JOURNAL		Nature 409 (6821), 685-690 (2001)
MEDLINE		21085660
PUBMED		11217851
REFERENCE	5	The FANTOM Consortium and the RIKEN Genome Exploration Research
AUTHORS		Group Phase I & II Team.
TITLE		Analysis of the mouse transcriptome based on functional annotation
JOURNAL		of 60,770 full-length cDNAs
MEDLINE		Nature 420, 563-573 (2002)
PUBMED		6 (bases 1 to 4485)
REFERENCE		Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P.,
AUTHORS		Fukuda, S., Furuno, M., Hanehara, T., Hara, A., Hashizume, W.,
		Hayashida, K., Hayatsu, N., Hizumoto, K., Higoka, T., Hirozane, T.,
		Horii, F., Imocani, K., Ishii, Y., Itch, M., Kagawa, I., Kasukawa, T.,
		Katoh, H., Kawai, J., Kojima, Y., Kondo, S., Kono, H., Kouda, M.,
		Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M.,
		Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohtsuo, N.,
		Okazaki, Y., Salto, R., Salto, H., Saitoh, C., Sakai, K., Sakazume, N.,
		Sano, H., Sasaki, D., Shibata, K., Shingawa, A., Shiraki, T.,
		Sogabe, Y., Tagami, M., Tagawa, K., Takahashi, F., Takaku-Akahira, S.,
		Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A.,
		Muramatsu, M. and Hayashizaki, Y.
TITLE		Direct Submission
JOURNAL		Submitted (16-JUL-2001) Yoshinide Hayashizaki, The Institute of

COMMENT  
 CDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.  
 Please visit our web site for further details.  
 URL: <http://genome.gsc.riken.go.jp/>  
 URL: <http://fantom.gsc.riken.go.jp/>.  
 Location/Qualifiers  
 1. 4485  
 /organism="Mus musculus"  
 /mol\_type="mRNA"  
 /strain="C57BL/6J"  
 /db\_xref="PANTOM\_DB:4831440K17"  
 /db\_xref="taxon:10090"  
 /clone\_id="4831440K17"  
 /cruise\_type="head"  
 /cruise\_lib="RIKEN full-length enriched mouse cDNA library"  
 /dev\_stage="0 day neonate"  
 112. 4439  
 /note="putative  
 weakly similar to alpha-2-macroglobulin (fragments)  
 [limulus polyphemus] (PIR|A36260, evidence: FASTV,  
 51.8kID: 76.7%length, match=168)"

BASE COUNT 1277 a 970 c 1010 g 1228 t

ORIGIN

Alignment Scores:  
 Pred. No.: 0 Length: 4485  
 Score: 5343.00 Matches: 1036  
 Percent Similarity: 83.70% Conservative: 176  
 Best Local Similarity: 71.55% Mismatches: 206  
 Query Match: 72.71% Indels: 31  
 DB: 11 Gaps: 7

US-10-020-095-4 (1-1428) x AK029247 (1-4485)

```

QY      1 MetGngIyPProProLeuLeuThraAlaHisLeuLeuCyValCyThraAlaLeu 20
DB      112 ATGCCAGATCGGAGGCTCCTCGAGCCGCGCCACTTCTCTGCTTGGCCGCGCTG 171
QY      21 AlaValAlaPProGlyProArgPheLeuValThraAlaProGlyVilIleIlaPProGly 40
DB      172 GCC--GGCCCGAGGTCCCGCTTGTGGAGACGCCCGAGGATCTCAAGACCCGAGCA 228
QY      41 AenValThrIleGlyValGAluLeuLeuGluHisCyProSerGlnValThraValAla 60
DB      229 AATGAGCATATTGGGTGGATCTCTCGAAATATACCCCCACAGAGTCCTTGAAGGCT 288
QY      61 GluLeuLeuLeuThraLseSerAnLeuThraLseValLeuGluAlaGluGlyAlaPhe 80
DB      289 CAGGTGTTCAAGATATAGCTTCCAAACAATCGAGATCATCTCGAAGCAAGAAGACTTT 348
QY      81 GluLySGlySerPheLyGThreuthrLeuThreuproSerLeuProLeuAnSerAlaApo 100
DB      349 CACAGAGGCCATTTCAGACTCTCGTTCTTCGGAGCATACCTTGACGTGCAGATAG 408
QY      101 IleyrGluLeuArgValThrcGlyrGThrglnaPgluIleLeuPheSerAenseTr 120
DB      409 ATTATGAGCTACACATTAACGCAATTCAGGAATGAGATGATATTCTCCAAACAGACA 468
QY      121 ArgLeuSerPheGluThrLyArgIleSerValPheIleGlnThraApLyAlaLeuTy 140
DB      469 CGTTTAACATTGAGAGCAAGAGTATATCTGCTTCATTCACAGACAGAGGACTTAC 528
QY      141 LysPProLyGlnGluValLysPheArgIleValThrLeuPheSerAspPheLyProTy 160

```

QY	515	ThrmctPheSerLeuThrProGluAsnSerTrpHisValIleValTyr	534
Db	1659	AGAACTTCTCTTAAACCAAGAGCTCTGGGCTCCAAAGGCGCTGATATATGCAAT	1718
QY	535	TyrIleGluAspArgGlyGluIleIleSerAspValLeuValIleProValIleValTyr	554
Db	1719	TATATGACAGAGAGTGGGAAATTTATAATGATTTCTTAATAAATCCCGCTTACACTTCT	1778
QY	555	PheIysAsnIleValLeuTyrTrpSerIysValIleValGluProSerGluIysVal	574
Db	1779	TTTGAATAATAGTAAAGAGCTGTTTGGAGTAAACCTACAGTCAAGCCATCGATTAAGGTC	1838
QY	575	SerLeuAspGlyIleSerValThrGlnProAspSerIleValGlyIleValAlaValAspIys	594
Db	1839	TCTCTCAGAGATCTCTGCACACAGCTGTGACCTCCCTGGGGGATTTGAAGCTGTTCACAAA	1898
QY	595	SerValaLeuMetCAsnAlaSerAspAspIleThrMetGluAsnValIleIleGluLeu	614
Db	1899	AGGTAACTTGAATGGAAACTCCAAACGATTTTCAATTGMAACCAATGGTCCATGAACCTG	1958
QY	615	GluLeuTyrAsnThrGlyTyrTyrLeuGlyMetPheMetAsnSerPheAlaValPheGln	634
Db	1959	GAACTTTTAAACACAGATATTAATCTAGGATGTTTCAATCTCTTCGCTGATTTTACG	2018
QY	635	GluCysGlyLeuThrPvalLeuThrAspAlaAsnLeuThryIleAspGlyVal	654
Db	2019	GAATCTGCTCTCTGGGTGTGACAGATGCCAACCCTTTAAGACGATGACATTTGAAGAAGTC	2078
QY	655	TyrAspAsnAlaGluTyrAlaGluArgPheMetGluGluAsnGluGlyHisIleValAsp	674
Db	2079	TACGATCTGAAGAGTATCTGAAACGCTTGCAAGAAATGAGGCAAACTGGTGTAT	2138
QY	675	IleHisAspPheSerLeuGlySerSerProHisValArgIleHisPheProGluThrTrp	694
Db	2139	TTTGAAGATGCTCTTCAAGTTACAACTACATGTCCAGGAAGATTTTCCAGAAACATCG	2198
QY	695	IleTrpLeuAspThrAsnMetGlyTyrArgIleTyrGlnGluPheGluValThrValPro	714
Db	2199	ATTGGCTAGACGCGCTACATGGAGATCTTAATGAAATTTATGAAGATTTGAAGTACTGTACT	2258
QY	715	AspSerIleThrSerTrpValAlaThrGlyPheValIleSerGluAspLeuGlyLeuGly	734
Db	2259	GACTTCACACACTTCTTGGGTGGCTTCGGCTTTTGTCACTCAGAAAGACCTGGGTTTGGG	2318
QY	735	LeuThrThrThrProValGluLeuGlnAlaPheGlnProPhePheIlePheLeuAsnLeu	754
Db	2319	TTTAACACCGTTCCAGACGAGCTGACGCTTCCAAACCGTTTTCCTTTCTCTAACTT	2378
QY	755	ProTyrSerValIleArgGlyGluGluPheAlaLeuGluIleThrIlePheAsnTyrLeu	774
Db	2379	CGTAACTGTATTATCAAGAGGTGAAGAGTTCCTTTGGAAATATCAATCGTCAATTATTTG	2438
QY	775	LysAspAlaThrGluValIleIleIleGlySerAspIysPheAspIleLeuMet	794
Db	2439	AAAGTACAAATTAAGTGTGATCTCATGAGAAAGTGCACACTTGTGATTATTTAATG	2498
QY	795	ThrSerSerGluIleAsnAlaThrGlyHisGlnIleThrLeuLeuValProSerGluAsp	814
Db	2499	ACTTAAATGACCCCAATGGACCAATATACCGGAAGACTGTTCAGATTCCAGGGAGTAT	2558
QY	815	GlyAlaThrValLeuPheProIleArgProThrHisLeuGlyGluIleProIleThrVal	834
Db	2559	GGGGTCACTCTGTTTTCCTCAAGCCACGACTTTGGAGAGATTTCCATACACTG	2618
QY	835	ThrAlaLeuSerProThrAlaSerAspAlaValThrGlnMetIleLeuValIysAlaGlu	854
Db	2619	ACGGCTGCTCACCCACTGCTCTGATGCTGTCAACCAACAATTGATGAACCTCGAA	2678
QY	855	GlyIleGlyIysSerTyrSerGlnSerIleLeuLeuAspLeuThrAspAsnArgLeuGln	874
Db	2679	GGGATAGAAAGTGGTATTTGCAATCTGTATTATGATCTGACAGATAGCAACGTTGAA	2738

[illegible]

Db	3816	CTTGATGCGCTAGATGCCATCACCGCTTAATGTTTGGCATGGCTGGGATTTGCAATT	1257
Qy	1238	CysGlnLeuAenValValTYrAsnValAlaSerGlySerSerAgaGArGArSer	1257
Db	3876	TGCCAGCTTAATGTTGACTTAACGTGAAGAGTTTCAGGTTCTTCTAAMAAACGGAGATCT	3395
Qy	1258	ILeIInGngInGluAlaPheAspLeuAspValAlaValysGluAenLysAspAspLeu	1277
Db	3936	ACCAAAATCAAGAAAGTTTGTGATTAGAGTCATTTGTG--AAATAATGAGACGACATTT	3992
Qy	1278	AsnHisValAspLeuAenValCysThrSerPheSerGlyProGlyAaSerGlyMetAla	1297
Db	3993	AGTACCTCGAATCTGAATGTGTGTGCAAGTCACTGTGGTTCAAGAGACAGCATGTGTC	4052
Qy	1298	LeuMetGluValAsnLeuLeuSerGlyPheMetValProSerGluAlaIleSerLeuSer	1317
Db	4053	CTTATGGAAGTGAACCTTCTCAGTGGCTTTAGTGCATCTTCAGATCAATTCCTCTAGT	4112
Qy	1318	GluThrValIlyIlyValGluTYrAspHisGlyIlyLeuAenLeuTYrLeuAspSerVal	1337
Db	4113	GAGACCCCTGAAGAAAGTGAATAAGATACAGGAAACCTTAACCTTATTTAGATCTGTA	4172
Qy	1338	AsnGluThrGlnPheCysValAsnIleProAlaValArgAsnPheLysValSerAsnThr	1357
Db	4173	AATGAATCCAGTTTGTGTTAATATTCCTCAGTGTGAGACTACAAAGTTTCAATATTT	4232
Qy	1358	GlnAspAlaSerValSerIleValAspTYrTYrGluProArgArgGlnAlaValArgSer	1377
Db	4233	CGAGATGGTTCAGTATCTGTATATGACTACTATGAAACAAAGAGACAGGAGTGCAGAC	4292
Qy	1378	TYrAsnSerGluValIlyIlyLeuSerSerCysAspLeuCysSerAspValGlnGlyCysArg	1397
Db	4293	TACAACCAACCCAGGAGGAGTGTCTCTCAATGTACTCTCAGTCAGACACC--AACTGCAAG	4349
Qy	1398	ProCysGluAspGlyAlaSerGlySerHisHisSerSerValIlePheIlePheCys	1417
Db	4350	TCACACACGGACGGAGCGACGAGACTCCTTCGACGTTCTTCGAGCCTTCGTCTTTGTC	4409
Qy	1418	PheLysLeuLeuTYrPheMetGlu	1425
Db	4410	TCCGTCCTTCTGTACTTTGTGCA	4433
RESULT 2			
AK036799		4725 bp	mRNA
LOCUS			linear HTC 05-DEC-2002
DEFINITION		Mus musculus adult female vagina cDNA, RIKEN full-length enriched	
		library, clone,930012E15 product:weakly similar to	
		alpha-2-macroglobulin (fragmente) [limulus polyphemus], full insert	
		sequence.	
ACCESSION	AK036799		
VERSION	AK036799.1	GI:26085434	
KEYWORDS	HTC; CAP trapper.		
SOURCE	Mus musculus (house mouse)		
ORGANISM	Mus musculus		
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
REFERENCE			
AUTHORS	Carninci, P. and Hayashizaki, Y.		
TITLE	High-efficiency full-length cDNA cloning		
JOURNAL	Wch. Enzymol. 303, 19-44 (1999)		
MEDLINE	99279253		
PUBMED	10349636		
REFERENCE	2		
AUTHORS	Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,		
TITLE	Itch, M., Kono, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.		
JOURNAL	Normalization and subtraction of cap-trapper-selected cDNAs to		
MEDLINE	prepare full-length cDNA libraries for rapid discovery of new genes		
PUBMED	Genome Res. 10 (10), 1617-1630 (2000)		
REFERENCE	3		
AUTHORS	Shibata, K., Itch, M., Aizawa, K., Nagaoke, S., Sasaki, N., Carninci, P.,		



Komno, H., Akiyama, J., Nishi, K., Kitsuana, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, B., Watabiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.  
RIKEN integrated sequence analysis (RISA) system -384-format  
Genome Res. 10 (11), 1757-1771 (2000)

TITLE  
JOURNAL  
MEDLINE  
PUBMED  
REFERENCE  
AUTHORS

4 Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y., Arakawa, T., Hara, A., Fukunishi, Y., Komno, H., Adachi, J., Fukuda, S., Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamana, S., Saito, T., Okazaki, Y., Gojobori, T., Bono, H., Kasukawa, T., Saito, R., Kadota, K., Matsuda, H., Ashburner, M., Batilov, S., Casavant, T., Fleischmann, W., Gaasterland, T., Gissi, C., King, B., Kochiwa, H., Kuwaji, P., Lewis, S., Matsuo, Y., Staudli, F., Suzuki, R., Tomita, M., Quackenbush, J., Schriml, L. M., Staudli, F., Suzuki, R., Tomita, M., Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H., Baldarelli, R., Barsh, G., Blake, J., Boffelli, D., Bojunga, N., Carninci, P., de Bonaldo, M. F., Brownstein, M. J., Bult, C., Fletcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D., Holtman, M., Hume, D. A., Kamiya, M., Lee, N. H., Lyons, P., Marchionni, L., Mashima, J., Mazzarelli, J., Mombert, P., Nordone, P., Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H., Sato, K., Schonbach, C., Seya, T., Shibata, Y., Storch, K. F., Suzuki, H., Toyooka, K., Wang, K. H., Weitz, C., Whitaker, C., Wilming, L., Wyshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawai, H., Kohetsuki, S. and Hayashizaki, Y.  
Functional annotation of a full-length mouse cDNA collection  
Nature 409 (6821), 685-690 (2001)  
21085660  
11217851

TITLE  
JOURNAL  
MEDLINE  
PUBMED  
REFERENCE  
AUTHORS

5 The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.  
Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs  
Nature 420, 563-573 (2002)  
6 (bases 1 to 4725)

Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hayashizaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hirakawa, T., Hirozane, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Katoh, H., Kawai, J., Kojima, Y., Kondo, S., Komno, H., Kouda, M., Koyama, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N., Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shingawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Tanaka, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Yamamatsu, M. and Hayashizaki, Y.  
Direct Submission

TITLE  
JOURNAL  
COMMENT

Submitted (16-JUN-2001) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute, 1-7-22 Suenitro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gs.riken.go.jp, URL: http://genome.gsc.riken.go.jp, Tel: 81-45-503-9222, Fax: 81-45-503-9216)

CDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.

Please visit our web site for further details.

URL: http://genome.gsc.riken.go.jp/  
URL: http://fantom.gsc.riken.go.jp/  
Location/Qualifiers

FEATURES  
SOURCE

1. .4725  
/organism="Mus musculus"

misc\_feature  
/note="putative  
weakly similar to alpha-2-macroglobulin (fragments)  
[Limulus polyphemus] (PIR|A3260, evidence: FASTY,  
51.8%ID, 76.7%length, match=168)"  
4707..4712  
/note="putative"  
polyA\_signal  
4725  
polyA\_site  
/note="putative"  
/note="putative"  
BASE COUNT 1334 a 1016 c 1066 g 1309 t  
ORIGIN  
Alignment Scores:  
Pred. No.: 0 Length: 4725  
Score: 5304.00 \*Matches: 1037  
Percent Similarity: 83.49% Conservative: 177  
Best Local Similarity: 71.32% Mismatches: 202  
Query Match: 72.18% Indels: 38  
DB: 11 Gaps: 8  
US-10-020-095-4 (1-1428) x AK036799 (1-4725)  
QY 1 MetGInGlyProProLeuLeuThraAlaHisLeuLeuCyValCyThrAlaAlaLeu 20  
DB 97 ATGCCCAATCGAGGCTCTTAAGCGGCCCACTTCTCTGCTTGGCGCGCTG 156  
QY 21 AAlaValAlaProGlyProArgPheLeuValThraAlaProGlyIleIleArgProGlyGly 40  
DB 157 GCC---GCCCAAGGATCCCGCTTCTGCTGAGACGCCAGGAGATCATGACCCGAGCA 213  
QY 41 AsnValThrIleGlyValGluLeuLeuGluHisCySPProSerGlnValThyValysAla 60  
DB 214 AATGTGACTAATGGGTGGGTGATCTCTGGAATAAGCCCCCAAGGTCTGTAAAGCT 273  
QY 61 GluLeuLeuYsThrAlaSerAsnLeuThraValSerValLeuGluAlaGluGlyAlaPhe 80  
DB 274 CAGGTTCAGATGATGCTTCAACAATCGAGATCATCTGGAAGCAGAGAGGTCTT 333  
QY 81 GluYsGlySerPheLeuThraLeuThraProSerLeuProLeuAsnSerAlaAspGlu 100  
DB 334 CACAGAGCCCAATTCATCAACTCTCTTCCGACATCACTCTGACAGATGACATAG 393  
QY 101 IleYrGluLeuArgValThrGlyArgThrGlnAspGluIleLeuPheSerAsnSerThr 120  
DB 394 ATTATGAGCTACATCAATCAAGCAATCAAGATGATGATGATCTCCAAACAGCACA 453  
QY 121 ArgLeuSerPheGluThraValArgIleSerValPheIleGlnThrAspValAlaLeuThr 140  
DB 454 CGTTAAACATTTGAGAGCAAGATGATATCTGCTCATTCAGACAGACAGACATTTTAC 513  
QY 141 LysProLysGlnGluValIlePheArgIleValThraLeuPheSerAspPheLeuProGly 160  
DB 514 AAGCGAAGCAGAGAGTAAATTTCCGAGTCTGACACTCTGCTGATCTGATCAAGCTTAT 573  
QY 161 LysThrSerLeuAsnIleLeuIleLysAspProLysSerAsnLeuIleGlnGlnThrLeu 180  
DB 574 AGGACATCCGTCGACATTTTCAATGAGATCCGAAGTATGATGATCAACAGAGGTTT 633  
QY 181 SerGlnGlnSerAspLeuGlyValIleSerYsThrPheGlnLeuSerSerHisProIle 200  
DB 634 TCTCAGAAAGGTGATCTTGAGTGTTCACAACTTTCATCACTCTTAATCAATCA 693  
QY 201 LeuGlyAspTPSerIleGlnValGlnValAsnAspGlnThrYrYrGlnSerPheGln 220

Db 694 TTGGTGACTGGTTCATTGAGTTTCAAGTATGATGACGAATATTAACGATCATTTTCAG 753  
 Qy ValSerGluTyrValLeuProLeuPheGluValThrLeuGlnThrProLeuTyrCysSer 240  
 Db 754 GTCTTAAGATATGCTTATCCAAAGTTGGAAGTCACTGCGACAGACCCGCTGTATTTGCTCC 813  
 Qy MetAsnSerIleuAsnGlyThrIleThrAlaIleGlyTyrGlyIleProVal 260  
 Db 814 TTGAAGTCTAAACAGTTAAATGCTCCGTCATAGCAAGTATATCATACGGGAAGCCAGT 873  
 Qy 260 LysGly-AspValThrLeuThrPheLeuProLeuSerPheTrp-GlyValLeuAsn 279  
 Db 874 GAAGGAGAGCTGTCTCATCATATTTTACCTTTGCTTTTGGGAGAAAGAAAGAAAT 933  
 Qy 280 IleThrIleThrPheIleAsnGlySerAlaAsnSerPheAsnAspGluMet 299  
 Db 934 ATCAAAAAGTTTGGAGATTAATGATTTGCAAACTTCTTTTGTATTAATTAAGATG 993  
 Qy 300 LysAsnValMet-----AspSerSerAsnGlyLeuSerGluTyrLeu 313  
 Db 994 AAAAAGGTAATGATTTGAAACCGCTAACGATGTCTCGAAGGAGATTAAGAAATGTG 1053  
 Qy 314 AsnLeuSerSerProGlyProValGluIleLeuThrThrValThrGluSerValThrGly 333  
 Db 1054 GACCCGTCCTTCTCGGCGCGGACAGATTAATGACTCTGACGAGATCACTCACAGGT 1113  
 Qy 334 IleSerIleAsnValSerThrAsnValPhePheGluGlnIleAspTyrIleLeuPhe 353  
 Db 1114 ATCTCAAGATGGACAGACCAATGTGTTTTCAGACACATGATTAATCATCATTTGAGATT 1173  
 Qy 354 PheAspTyrThrThrValLeuLysProSerLeuAsnPheThrAlaThrValLysValThr 373  
 Db 1174 TTTCATTAATAGACTGTCTTGAAACCATCTCTCAACTTCACAGCCAGCTGGAAGTCAAG 1233  
 Qy 374 ArgAlaAspGlyAsnGlnLeuThrLeuGluValArgAsnAsnValIleThr-Val 393  
 Db 1234 CGCTCATGTGCAATCACTAACTCCGAGAAATAGAAATGATTTGGTCAACAGTAC 1293  
 Qy 393 LThrGluArg-----AsnTyrThrGluTyrTrpSerGlySerAsnSerGlyAsnGln 411  
 Db 1294 GACCCAGCGGAAAAATTAACCACTGAG-----AGTCAGAGAGATTCAGGA 1338  
 Qy 411 SmetGluAlaValGlnLysIleAsnTyrThrValProGlnSerGlyThrPheLysIleG 431  
 Db 1339 AATGATTAATATCCAGACCCGTAATTAATATCATCCCAAAATGGAATCAATTAAGATTGA 1398  
 Qy 431 uPheProIleLeuGluAspSerSerGluLeuGlnLysAlaTyrPheLeuGlySer 451  
 Db 1399 ATTCCTCAGTCAATGTCCATTTCCGCGGAGCTAACCTAAGGCTTATTTTCTGTGATGCTAC 1458  
 Qy 451 sSerSerMetAlaValIleSerLeuPheLysSerProSerIleThrTyrIleGlnLeu 471  
 Db 1459 AAGTTCGTGACAGTCCACATATGTTCACTCTCCATAGTAAGCGATACCACTGCTTA 1518  
 Qy 471 sThrArgAspGluAsnIleLysValGlySerProPheGluValLeuValIleSerGlyAsn 491  
 Db 1519 AACCAAGATGAATATTAAGTGGGCTACCTTTGATTTGATGTGTGTCACACAG 1578  
 Qy 491 sArgLeuLysGluLeuSerTyrMetValIleSerArgGlyGlnLeuValAlaValGly 511  
 Db 1579 ACAATTCAGAGCTTAAGCTAATATGATTAATCAAGGACAGTGTGCTGACGACCA 1638  
 Qy 511 sGlnAsnSerThrMetPheSerLeuThrProGluAsnSerTrpThrProLysAlaCysVal 531  
 Db 1639 ACAAAAGTCAAGAACTTCTTTTAAACACAGAGCTTCTCGGCGCTCCAAAGGCTGTAT 1698  
 Qy 531 IleValTyrTyrIleGluAspArgGlyGlnIleIleSerAspValLeuLysIleProVal 551  
 Db 1699 AATCGCATATTAATTTGACAGAGATGGGAAATTAATTAAGATATTTTAAAAATCCCGCT 1758  
 Qy 551 LglnLeuValPheLysAsnLysIleLysLeuTyrTrpSerIleValLysAlaGluProSe 571  
 Db 1759 TCAGCTTGTTTTGAATAATAGTAAAGCTGTTTGGAGTAAACCTACAGTCAAGCATC 1818

Qy 571 rGluLysValSerLeuArgIleSerValThrGlnProAspSerIleValGlyIleValAl 591  
 Db 1819 TGATTAAGTGTCTCTCAGAGATCTCTGACACAGTGTGACTCCCTGGTGGATTTGAGC 1878  
 Qy 591 aValAspLysSerValAsnLeuMetAsnAlaSerAsnAspIleThrMetGluAsnVal 611  
 Db 1879 TGTGTGCAAAAGTGAATGATTAAGTGAACCTCCAAAGATTAATCAATGAAACCATGTGT 1938  
 Qy 611 LHisGluLeuGluLeuTyrAsnThrGlyTyrTyrIleGluGlyMetPheMetAsnSerPheAl 631  
 Db 1939 CCAATGAACCTGAACCTTTATTAACACAGAAATTAATCTAGGCAATTTATGAACTCTTCGC 1998  
 Qy 631 aValPheGlnGluCysGlyLeuTyrValLeuThrAspAlaAsnLeuThrLysAspTyrIle 651  
 Db 1999 TGTATTTCAGGAATGTGTGTCTCGGCTGTGACAGATGGAACCTTATTAAGATATGAT 2058  
 Qy 651 eAspGlyValTyrAspAsnAlaGluTyrAlaGluArgPheMetGluGluAsnGluLys 671  
 Db 2059 TGATGAAGTCTACGATACGAAAGATATTCTGAACTTGTGACAGGAAATGAGCCAA 2118  
 Qy 671 sIleValAspIleHisAspPheSerLeuGlySerSerProHisValArgLysHisPhePr 691  
 Db 2119 CTGTGTGATTTGAAAGATGCTTCTCAGTTAACAAATGATGATGACGAAAGAAATTTCC 2178  
 Qy 691 oGluThrTrpIleTrpLeuAspThrAsnMetGlyTyrArgIleTyrGlnGluPheGluVal 711  
 Db 2179 AGAAACATGATTTGGCTAGACCCCTAACATGGGATCATAGATTAATGAAGAAATTTGAGT 2238  
 Qy 711 LThrValProAspSerIleThrSerTyrValAlaThrGlyPheValIleSerGluAsp 731  
 Db 2239 TACTGTACCTGATCATCATCTTCTGGGCTGCTTGGCTTGTGATCATCAAGAACCT 2298  
 Qy 731 uGlyLeuGlyLeuThrThrProValGluLeuGlnAlaPheGlnProPheIlePr 751  
 Db 2299 GGGTTTGGGTAAACAACGTTCCAGACAGAGCTGCAAGCTTCAACCGTTTCTCTTT 2358  
 Qy 751 eLeuAsnLeuProTyrSerValIleArgGlyGluGluPheAlaLeuGluIleThrIlePr 771  
 Db 2359 CTGATCTTCCCTACTCTGTTATCAGAGTGAAGGTTTCTTGGAGTATCAATCGT 2418  
 Qy 771 eAsnTyrLeuLysAspAlaThrGluValLysIleIleGluLysSerAspLysPheAl 791  
 Db 2419 CAATTAATGAAGATACAAATTAAGCTTGATGATCATCATGAGAAAGTCAACAGCTTTGA 2478  
 Qy 791 rIleLeuMetThrSerSerGluIleAsnAlaThrGlyHisGlnGlnThrLeuValPr 811  
 Db 2479 TATTTTAATGACTTCAAAATGACCAATGACACATATACCGGAACACTGTTCAAGTTCC 2538  
 Qy 811 oSerGluAspGlyAlaThrValLeuPheProIleArgProThrHisLeuGlyGluIlePr 831  
 Db 2539 CAGGATTAATGGGCTGACTTGTCTTTCCATCCAGACCCACGACCATTTGGGAAGATTC 2598  
 Qy 831 oIleThrValThrAlaLeuSerProThrAlaSerAspAlaValThrGlnMetIleLeuVal 851  
 Db 2599 CATCACAGTGAACGCGCTCACCACCTGCTGTGATGCTGTCACACGACCAATTTGAGT 2658  
 Qy 851 LysAlaGluGlyIleGluLysSerTyrSerGlnSerIleLeuLeuAspLeuThrAsp 871  
 Db 2659 AAAGCTGAAGGATGAAAAGTGAATTCGAAATGTCTTAATGATTCGACAAATGAG 2718  
 Qy 871 nArgLeuGlnSerThrLeuLeuSerPheSerPheProProAsnThrValThrGln 891  
 Db 2719 CAACGTAAAGTAAGACAGCAATTAATGATTTCTCTTCCCTCGATACGTCATGCG 2778  
 Qy 891 ySerGluArgValGlnIleThrAlaIleGluAspValLeuGlyProSerIleAsnGlyLe 911  
 Db 2779 CAGTGAAGAGTTCAGATCAACAGCAATGGAAGTATCTTGTGTTCTCCATCAATGAGCTT 2838  
 Qy 911 uAlaSerLeuIleArgMetProTyrGlyCysGlyGluGlnAsnMetIleAsnPheAlaPr 931  
 Db 2839 ATCTTCACGATCCGATGCTTATGATGTGTGAAACAGAAATGATATATTTGCTCC 2898

```

QY 931 cAsnIeTyrlleLeuAspTyrlleuThrlslyslglnleuThrAspAsnleuLysgl 951
Db 2899 AAATATTACATTTGGATTATCTGACTTAACAGAAACAGCTGACAGTTAATTTAAAGA 2958
QY 951 uLysAlaLeuSerPheMetArgLglnLYTyrGlnArgGluLeuLysGlnArgGluAs 971
Db 2959 AAGAGCCCTTTCATATATAGAGGAGGATTAACAAAGGAGGCTTCTCTATCAAGAGGAAACA 3018
QY 971 pGlySerPheSerAlaPheGlyAsnTYrAspProSerGlySerThrlPheSerAlaPhe 991
Db 3019 TGGCTCTTTCAGTGGCTTTGGGAGACTTACCTTCTGGAGACACTGGCTCTGACACT 3078
QY 991 eValLeuArgCysPheLeuGlnAlaAspProTyrIleAspIleAspGlnAsnValLeuH 1011
Db 3079 TGTTTTAAGATGCTTTCTGGAGCTGATTAATATAGATATGATATGATAGAGATGTGTACA 3138
QY 1011 sArgThrTYrThrlPheLysGlnLYHlsGlnLysSerArgGlnLYGlnPheTyrAspProG 1031
Db 3139 CAGAACATATACCTTGGCTTAATGCACTTAAGAAATTCATATGATGTAATTTGGAGCCAG 3198
QY 1031 YArgValIleHisSerGlnLeuGlnLYGlnLYAsnLysSerProValThrlleuThrAlaTY 1051
Db 3199 AAGATGATTCACATGAACTTCAAGGTGACCAAAAGCCAGTATACCTTACCGCCTTA 3258
QY 1051 rIleValThrSerLeuLeuGlnLYTyrArgLysTYrGlnProAsnIleAspValGlnLys 1071
Db 3259 TATTTGACTTCTGCTGGAGTATCAAAAGATACAGCTTATATGATATGATATGATATGAT 3318
QY 1071 rIleHisPheLeuGlnSerGlnPheSerArgGlnLYIleSerAspAsnTYrThrlleuAla 1091
Db 3319 AATCAAGTTTGTGAAATTTGAATTCAGCAGAGAAATTCGAGCAATTTATACCTTGACAT 3378
QY 1091 uIleThrTYrAlaLeuSerSerValGlySerProLYsAlaLysGlnAlaLeuAsnMetle 1111
Db 3379 TATATCCATATGCTCTGCTCCACAGTGGGAGCCCTTAAGACAGAGAGGCTTATACTGCT 3438
QY 1111 uThrThrlPheArgLglnGlnLYGlnLYMetGlnPheTyrValSerSerGlnSerLYsle 1131
Db 3439 GATGCAGCGATCAGAAAGAGAGAGAGACACAGTTCTGTTGTCATCAGACCTGCGCT 3498
QY 1131 uSerAspSerTYrGlnProArgSerLeuAspIleGlnValAlaAlaTYrAlaLeuLeuSe 1151
Db 3499 CTCTGGGCTGTCGAGCGCGCTGCTGCTGACATCAAAATCCGCTCTTATGCACTCTGCG 3558
QY 1151 rHisPheLeuGlnPheGlnThrSerGlnLYrIleProIleMetArgTYrPheSerArgL 1171
Db 3559 GCACACTCTGCAC---CATGTCGTAGAGGAAATCCCGGTTATAGAGTGGCTCATCAGCA 3615
QY 1171 nArgAsnSerLeuGlnLYGlnLYPheAlaSerThrlGlnAspThrlThrValAlaLeuLYsAla 1191
Db 3616 AAGAAACAGCGCTGGAGGTTTGTATCCACTCAGGATCTGTGTGCGCCCTTAAGGCAAT 3675
QY 1191 uSerGlnPheAlaAlaLeuMetAsnThrlGlnArgThrAsnIleGlnValThrlThrl 1211
Db 3676 ATCTGAATTTTCAGCCCTTGTGCACAGAGAAATATACATATCAACGTACCGTCACAGG 3735
QY 1211 YProSerSerPro----- 1215
Db 3736 GCCCGGCATCCCAATCCATACACTCTCGAAATGATCTCCAGAACTCTTCTCTCTCA 3795
QY 1216 -----SerProLeuAlaValAlaGlnProMetAlaValAsnIleSerAlaAsnGlyPheG 1234
Db 3796 CCAGAAAGAGCTTATGCGCTAGATCCATACCGTTAATTTCTGCGCAAGCTCGG 3855
QY 1234 YPheAlaIleCysGlnLeuAsnValTYrAsnValLYsAlaSerGlySerSerArgAr 1254
Db 3856 ATTTCCTATTTCAGCTTAAGTTGACATATACGTGAAGGTTCTGAGGTTCTTCTTAAG 3915
QY 1254 GARGArgSerIleGlnAsnGlnAlaPheAspLeuAspValAlaValLYsGlnAsnLY 1274
Db 3916 ACGGAGATCTACCGAAATCAAGAGTTTATATATGACGTCATTTG---AATTAATCA 3972
QY 1274 sAspAspLeuAsnHisValAspLeuAsnValCysThrlSerPheSerGlyProGlyArgSe 1294

```

```

Db 3973 GACGACATTTAGTACCTCGAATCTGAATGTGTGCAACAGTCACTTGGCTTACAGAGGAC 4032
QY 1294 rGlyMetAlaLeuMetGlnValAsnLeuLeuSerGlyPheMetValProSerGlnAlaI 1314
Db 4033 AGGCAATGCTCTTATAGAGGAGAACCTTCTCAAGTGGCTTTAGTGCATCTTCAATTCAT 4092
QY 1314 eSerLeuSerGlnThrValLYsLYsValGlnLYTyrAspHisGlyLYsLeuAsnLeuTYle 1334
Db 4093 TCTCTGATGAGACCCCTGAAGAAAGTGAATATATATACGGAACCTCAACCTTATTT 4152
QY 1334 uAspSerValAsnGlnThrGlnPheCysValAsnIleProAlaValArgAsnPhelYsVa 1354
Db 4153 AGATCTGTAATATGATCCAGTTTGTGTAAATTCCTCCACTGAGAGACTACAAAGT 4212
QY 1354 lSerAsnThrlGlnAspAlaSerValSerIleValAspTYrTYrGlnProArgArgGlnAl 1374
Db 4213 TTCAATATTTGAGATGAGTTCAGTATCTGTATAGACTATATGAACAGAGACAGC 4272
QY 1374 aValArgSerTYrAsnSerGlnValLYsLeuSerSerCysAspLeuCysSerAspValG 1394
Db 4273 AGTGGAGAGCTTCAACACCCAGGTGACCTGCTCTGTTTACTTCACTCACTCAGACACC-- 4330
QY 1394 nGlyCysArgProCysGlnAspGlyAlaSerGlySerHisHisSerSerValIlePh 1414
Db 4331 -AAGTCAAGATCAGACAGCAGGAGCAGGACCTCCTTGAGAGCTTCTGAGGCTTCT 4389
QY 1414 eIlePheCysPheLYsLeuLeuTYrPheMetGln 1425
Db 4390 CGTCTTTGCTCCCTCTTCTGTACTTGTGCAA 4423

```

## RESULT 3

BC044072

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

Xenopus laevis (African clawed frog)

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;

Xenopodinae; Xenopus.

1 (bases 1 to 4641).

REFERENCE

AUTHORS

TITLE

JOURNAL

REMARK

COMMENT

Contact: XGC help desk

Email: cgapbs-remail.nih.gov

Tissue Procurement: Dr. Igor David

CDNA Library Preparation: Life Technologies, Inc.

DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)

http://www.systemsbio.org

contact: amadan@systemsbio.org

Anup Madan, Jessica Fahey, Erin Helton, Mark Kettelman, Anuradha

Madan, Stephanie Rodrigues, Amy Sanchez and Michelle Whiting

Clone distribution: MGC clone distribution information can be found

through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov

Series: IRAX Plate: 94 Row: b Column: 12

This clone was selected for full length sequencing because it

passed the following selection criteria: Hexamer frequency ORF

analysis, Similarity but not identity to protein

This clone has the following problem: frame shifted.

Location/Qualifiers

1..4641

/organism="Xenopus laevis"

/mol\_type="mRNA"

FEATURES

source

```

/db xref="taxon:8355"
/clone="IMAGE:4681494"
/issue_type="Embryo, stage 31/32, Xenopus"
/clone_id="NICHD_XGC_Emb4"
/lab_host="DH10B"
/notes="Vector: pCMV-SPORT6"
BASE COUNT      1322 a 1063 c 1147 g 1109 t
ORIGIN

Alignment Scores:
Pred. No.:      7e-147      Length:      4641
Score:          1480.00      Matches:      433
Percent Similarity: 46.30%      Conservative: 287
Best Local Similarity: 27.85%      Mismatches:   564
Query Match:     20.14%      Indels:       273
DB:              11          Gaps:          46

US-10-020-095-4 (1-1428) x BC044072 (1-4641)
Qy      49  LeuGluHicYbProserGlnValThrValLysAlaGluLeuLeuLysThrAlaSerAsn 68
      |||||
      226 CTGGAGGGGGGGAATGGAAGAAGCAAGATACAGCTCACCTCACAGATGGCCAAAGTAAC 285
Qy      69  LeuThrVal-----SerValLeuGluAlaGluValPheGluLysGlySerPheLys 86
      |||||
      286 ACAACTGTGTATGAGAAAGACATCAGACAGAAATCTTCACTTCAGATGTCTTCAATTGAG 345
Qy      87  ThrLeuThrLeuProserLeuProLeuAsnSerAlaAspGluLeuTyrgluLeuArgVal 106
      |||||
      346 -----GTGCGCGCT---CCCTCTGAAGAGATATGAAGAGGCGCCACCATGCGAGTT 333
Qy      107 ThrGlyArgThrGlnAspGluLeuLeuPheSerAsnSerThrArgLeuSerPheGluThr 126
      |||||
      394 TCATATCAGAGCGGAGAGAGACCATAT---AGCAAGCAGAGTAAGTTCTGGTGAAGAAA 450
Qy      127 LysArgLysSerValPheLeuGlnThrAspLysAlaLeuTyrglyLysProLysGlnGluVal 146
      |||||
      451 GCAGAAACCAAGTATCTGTTCAACAGACAAACCTTATATCAAAACAGAGACAGACAGTA 510
Qy      147 LysPheArgLysValThrLeuPheSerAspPheLysProTyrglyLysThrSerLeu---Asn 165
      |||||
      511 AGATTCCAGAGTTGTTCTTAAAGAAACCTCCACACAGAAAGATCAGGTGCCAAC 570
Qy      166 IleLeuLysAspProLysSerAsnLeuIleGlnGlnThrLeuSerGlnGlnSerAsp 185
      |||||
      571 ATGGAATACAGAGCCAGGAGAGAACCGTATGGCCAGTGGCTGAATGAGTCTCCAG 630
Qy      186 LeuGlyValLysLeuThrPheGlnLeuSerSerHisProIleLeuGlyAspTyrPser 205
      |||||
      631 CAAGGAATCCCTGACTCTCATTCCCTCTCTCCAGCCCTCCGTTGGGAGAAATATCC 690
Qy      206 IleGlnValGlnAlaAsnAspGlnThrTyrglyLysPheGlnValSerGluTyrglyVal 225
      |||||
      691 ATTCAATTA-----AAGAACATGTCATCTTCAAGTGTGGAAGAAATATGTT 738
Qy      226 LeuProLysPheGluValThrLeuGlnThrProLeuTyrglySerPheAsnSerLysHis 245
      |||||
      739 CTGCCCAAGATTGAGTCACTCTCCATCCCAAAAGTGTGTGCTGTCACATGCGCAA 798
Qy      246 LeuAsnGlyThrIleThrAlaLysTyrglyLysProValLysGlyAspValThr 265
      |||||
      799 TTCCTGTATAAATCTGTCCAGGTACATTATGAAACACAGTGCAGGGA----- 849
Qy      266 LeuThrPheLeuProLeuSerPheTyrGlyLysLysAsnIleThrLysThrPheLys 285
      |||||
      850 -----ACCTATTAAG 858
Qy      286 IleAsn---GlySerAlaAsnPheSerPhe-----AsnAspGluGluMet 299
      |||||
      859 GTGATGTGTGCGGAGCATTTACAGTACAGAAAGGCATTGGTATATGCAAAAGGAGA 918
Qy      300 LysAsnValMetAspSerSerAsnGlyLeu----- 309

```

```

Db      919 GATCTGTGCTGATTTTCACTGGGAAAGCTGATCACTGTGCTCCACATTCGATGTG 978
Qy      310 -----SerGluTyrglyLeuAspLeuSerSerProGlyProValGluLe-----LeuThr 325
      |||||
      979 AACTCTTCTTACTACATCTTCTATACAGCGAATGGCAATGAACATTGAGGGGGCTGGC 1038
Qy      326 ThrValThrGluSerValThrGlyLysSerArgAsnValSerThrAsnValPhePheLys 345
      |||||
      1039 TGGATTACAGAAAGATGGACAGGAATTCAGATTCTGCAACTGATTAATCCCAATTTCGA 1098
Qy      346 GlnHisAspTyrglyLeuGluPhePheAspTyrglyThrValLeuLysProSerLeuAsn 365
      |||||
      1099 TATGACCTCAACAGTGTCACTTCTCGATGCTGATGATTAATTAAAGCTGGGATTCGA 1158
Qy      366 PheThrAlaThrValLysValThrArgAlaAspGlyAsnGlnLeuThr----- 381
      |||||
      1159 TACATCGGAGCACTGAAGGTGTGATGTCAGTGAACCCCATCCCAACATCCAGTGTG 1218
Qy      382 ---LeuGluGluArgArgAsnAsnValValIleThr---ValThrGlnArgAsnTyrglyThr 399
      |||||
      1219 TATCTGATCGACACCGATGGAATAGTAACATGACCTGTATGACCATTAAGATGAGACTT 1278
Qy      400 GluTyrglyTyrPserGlySerLysSerGlyAsnGlnLysMetGluAlaValGlnLysIleAsn 419
      |||||
      1279 GCTCCCTTCAAGCTGACATACAGACAGATTGGAGCGA----- 1317
Qy      420 TyrglyThrValProGlnSerGlyThrPheLysIleGluPheProIleLeuGlu----- 436
      |||||
      1318 ---ACGGTGTCCCTGACCGGCAAGACACAGTGAAGAGACCTTTATGCAAGTACGTAC 1374
Qy      437 -----AspSerSerGluLeuGlnLeuLysValTyrglyPheLeuGly 449
      |||||
      1375 GGCATATGTTCCCTGAGCATGGGTGATGCGATCTTACACTCAACCGCTTTATATCCCGC 1434
Qy      450 SerLysSerSerMetAlaValHisSerLeu-----PheLysSerProSerLysThr 466
      |||||
      1435 AGCAAGAGTTTATTAAGCTCCACTCCAGAAATGGGAGTTCTGTGTAGGGGCAACAA 1494
Qy      467 TyrglyLeuLysThrArgAspGluAsnIleLysValGlySerProPheGluLeuVal 486
      |||||
      1495 GAAGTCCAGGTGAATATCTTATCAAGCCACATGAGTTGGAGAG----- 1539
Qy      487 ValSerGlyAsnLysArgLeuLysGluLeuSerTyrglyMetValValSerArgGlyGlnLeu 506
      |||||
      1540 -----GAGCAAAAGACATTGAGTACTACTGCGAGTTTCTAAAGTGCCATA 1590
Qy      507 -----ValAlaValGlyLysGlnAsnSerThrMet----- 516
      |||||
      1591 AAGAATCTGACGTTTAAAGATCCAAATTAATGATGTGAAGTCTTGAGAGCGGAG 1650
Qy      517 -----PheSerLeuThrProGluAsnSerTyrglyProLysAlaCysValIleValTyrgly 534
      |||||
      1651 GTTCCATCAAACTCAAGCTCACTGTGACGTCTCTCCACACTTCTGCAATTCGACTTAC 1710
Qy      535 TyrglyLeuAspAspGlyGluLeuIleSerAspValLeuLysIleProValGlnLeuVal 554
      |||||
      1711 GTTTTATCCCGTGTGAGAGATGTGTCAGATGCGCAAAAGTTCAAGCTACAAAGTGC 1770
Qy      555 PheLysAsnLysIleLysLeuTyrglyTyrPserLysValLysAlaGluProSerGluLysVal 574
      |||||
      1771 TTTAGAAATTAAGGTGACATGTGCTTCTCCAGATGAAGTCTTACAGATCAAGATGTT 1830
Qy      575 SerLeuAsnIleSerValThrGlnProAspSerIleValGlyIleValAlaValAspLys 594
      |||||
      1831 TCCCTTCAAGTTCAGGCACT---GCCGGGTCTTCTGTGCTGCTGAGAGTGTGATTAAC 1887
Qy      595 SerValAsnLysMetAsnAlaSerAsnAspIleThrMetGluAsnVal----- 610
      |||||
      1888 AGTGTGTATTAATGAACCTGACCAAGACGCGGCAAGATTCGGGAGCTCTTC 1947
Qy      611 -----ValHisGluLeuGlu----- 615
      |||||
      1948 CCAGTCAATGATTTTGAATTAATGATTAATCGATTCAAGAAAGCTGAGATCACTGTCCA 2007

```

QY 616 -----LeuTyraAsnThr 619  
 Db 2008 TTCGACCCCTTCTTCTTGCGACCATTCATAGAAATTTTGAATTCATTATATATATCC 2067  
 QY 620 GlyTyrTyrLeuGlyMetPheMetAsnSerPheAlaValPheGlnGluCysGlyLeuTrp 639  
 Db 2068 CCATGGCAATATCCAGAGAACGTATGTCTACGGCTCTCTTCAAGAAAGTGCACCTAATAA 2127  
 QY 640 ValLeuThrAspAlaAsnLeuThrTyrAspTyrIleAspGlyValTyrAsp----- 656  
 Db 2128 ATCATCACCAGCGCTGCACATTAGAAACCATGGTGTGTCAACATTATGATTTTTTTAAG 2187  
 QY 657 -----AsnAlaGluTyrAlaGluThrPheMetGluGlnAsnGluYHsIleValAsp 674  
 Db 2188 GCCAGGGCTACAGAGCTTCAGAGCATCTGAGACCGAATGATTC-ATTGGTTTCC 2246  
 QY 675 IleHisAspPheSerLeuGlySerSerProHis----- 685  
 Db 2247 CTGGGATA-TCTGCTGCCGATCTTATACCAACAGCTACAGAACAGAGTCTGCAGCC 2305  
 QY 686 -----ValArgLysHisPheProGluThrTrpIleTrpLeuAspThrAsnMetGly 702  
 Db 2306 GTTGAATATATCCGAACCTACTTCTCAGAGCATGGAATATGGAGCTGGCATCTGTGGGC 2365  
 QY 703 TyrArgIleTyrGlnGluPheGluValThrValProAspSerIleThrSerTrpValAla 722  
 Db 2366 GAGTCTGGTACAGAGAGCTCATCGCGAGTCCAGACACCATACAGACTCGAATGCT 2425  
 QY 723 ThrGlyPheValIleSerGluAspLeuGlyLeuThrThrProValGluLeu 742  
 Db 2426 GGAGCTTCTGTATGGAGCCCACTG---GGATTGGGATC---TCTTCAACCACTCCCT 2479  
 QY 743 GlnAlaPheGlnProPhePheIlePheLeuAsnLeuProTyrSerValIleArgIleGlu 762  
 Db 2480 CGAGTCTTCAGCCCTCTTGTGTGAGAGCTGACTGCGTACTGTGTGGTGGAGAGAG 2539  
 QY 763 GluPheAlaLeuGluIleThrIlePheAsnTyrLeuLysAspAlaThrGluValIleVal 782  
 Db 2540 TCATTGCTCTCAAGGCACTCTTAACTACCTGAAACAGCTGATTAAGTTTACAGAC 2599  
 QY 783 IleIleGluLysSerAspLysPheAspIleLeuMetThrSerSerGluIleAsnAlaThr 802  
 Db 2600 GCCCTGGAGCCCTTCAAGAGTTGGAG-----GAGGAACCATGTCT 2641  
 QY 803 GlyHisGlnGlnThrLeuLeuValProSerGluAspGlyAlaThrValLeuPheProIle 822  
 Db 2642 GACTGTCTAGTATAGAGCTGTGTCTGTGACAGAGAGACCAAGACCTTCTACTGAACT 2701  
 QY 823 ArgProThrHisLeuGlyGluIleProIleThrValThrAlaLeuSer----- 838  
 Db 2702 AAAGCTTCTAACCTGGAGAGGTGAATGTCAACGTAAGACGCTGTGATCTCAG 2761  
 QY 839 -----ProThrAlaSerAspAlaValThr 846  
 Db 2762 GACTTGTGCATATATGAATCCCTATAGTCCCTTAAACAGGGAGATTCAGACCTGTGAT 2821  
 QY 847 GlnMetIleLeuValIleAlaGluIle-----GluLysSerTyrSerGlnSerIle 864  
 Db 2822 AAGCACAATCTGTGACAGCGAGGTCTGTGAAGAAAAGCTCACAGTTCTCTGCTC 2881  
 QY 865 LeuLeuAspLeuThrAspAsnArgLeuGlnSerThrLeuLysThrLeuSerPheSerPhe 884  
 Db 2882 TGCTGCCAAGCTGTGATGATATCTTAAACA-----GAGGAATATCTCTTAAGTCC 2935  
 QY 885 ProProAsnThrValThrGlySerGluArgValGlnIleThrAlaIleGlyAspValLeu 904  
 Db 2936 CCGTGAATATATCTTGAAGATTCAGAAAGACATATGTGACGCTCTGAGATCTGATG 2995  
 QY 905 GlyProSerIleAsnGlyLeuAlaSerLeuIleArgMetProTyrGlyCysGlyGluGln 924  
 Db 2996 GGAACAGCATGCAAGAACCTGACCGCTCTCTGGCCATGCGTATGATGTGGGAGCAG 3055

QY 925 AsnMetIleAsnPheAlaProAsnIleTyrIleLeuAspTyrLeuThrLysLeuGln 944  
 Db 3056 AACATGGTTCTCTTGTGCCCCCAATATCTTCAATCTTCAAGTATCTGAAAAGACATCACC 3115  
 QY 945 LeuThrAspAsnLeuLysGlyValAlaLeuSerPheMetArgGlnGlyTyrGlnArgGlu 964  
 Db 3116 CTGACCCCTGAGATCCAGAGCAAGCCAAAGGTTCTTGAAGCGGGCTCCAGAGCAG 3175  
 QY 965 LeuLeuTyrGlnAlaGluAspGlySerPheSerAlaPheGlyAsnTyrAspProSerGly 984  
 Db 3176 CTGACATCAACACGGATGACGATCATACAGCCGATTTGGGCAAGAGCAAGAGGGG 3235  
 QY 985 SerThrTrpLeuSerAlaPheValLeuArgCysPheLeuGlnAlaAspProTyrIleAsp 1004  
 Db 3236 AACACCTGCTGACTGATCTTGTGTGAAGTCTTTCAGCAAGGCTCGCTCATACATCTTC 3295  
 QY 1005 IleAspGlnAsnValLeuHisArgThrTyrThrTrpLeuLysGlyHsGlnLysSerAsn 1024  
 Db 3296 ATAGATGAAGCTCACTAATAACATTTCTTCACTTGGCTGAAGAACATCAACATGAAGT 3355  
 QY 1025 GlyLysPheTrpAspProGlyArgValIleHisSerGluLeuGlnGlyValAsnLysSer 1044  
 Db 3356 GCGTGTTCAGAGTGTGGGAGACTTTCATATATGCTATGAAGGAGGAATGATGAT 3415  
 QY 1045 ProValThrLeuThrAlaTyrIleValThrSerLeuGluGlyTyrArgLysTyrGlnPro 1064  
 Db 3416 GAACCTCTCTCTGCTGTATATGTAACAATGGCTTTGCTA-----GAACCT 3460  
 QY 1065 AsnIleAspValGlnGluSerIleHisPheLeuGlnSerGluPheSer----- 1080  
 Db 3461 GGTCTATCTGTACAGAGCCCAATG-----GTGAGGGTGTCTTATCTGCTGAGAGAG 3514  
 QY 1081 -----ArgGlyIleSerAspAsnTyrThrLeuAlaLeuIleThrTyrAlaLeuSerSer 1098  
 Db 3515 GCTGCCAAAGATGTAGAGATATCTTACCCAGAGCTCTGCTGCGCTACACTTCACTTCG 3574  
 QY 1099 ValGlySerProLysAlaLysGlu---AlaLeuAsnMetLeuThrTrpArgAlaGluGln 1117  
 Db 3575 AGTGTGATATGTAACACTCAGGAGATGCTGTGGCTAACTGATGAGAAGCTGTATAA 3634  
 QY 1118 GluGlyGlyMetGlnPheTrp-----ValSerSerGluSerLysLeuSerAsp 1133  
 Db 3635 AGTGAAGGACAGTTACACTGCGGAAGAAACCTCGCAGAGCTCCTGTGATTAACCTTAC 3694  
 QY 1134 SerTrpGlnProArgSerLeuAspIleGluValAlaAlaTyrAlaLeuLeuSerHisPhe 1153  
 Db 3695 TGTATCCAGGCTCCCTCAGCGAGGTGAGTTAATCTCTATGTATCTTGAACCTTCTC 3754  
 QY 1154 -----LeuGlnPheGlnThrSerGluGlyIleProIleMetArgTrpLeuSerArgGln 1171  
 Db 3755 TCTGGCCTCAAAAAGAGCTGGGAGAACTTCTGAGTATGATCAACTGGCTAGACAGCAG 3814  
 QY 1172 ArgAsnSerLeuGlyGlyPheAlaSerThrGlnAspThrThrValAlaLeuLysAlaLeu 1191  
 Db 3815 CAATATCATATGGGGATCTCTTCTTACTCAGACACAGGTTGTAGCTTCAAGGCTCTT 3874  
 QY 1192 SerGluPheAlaAlaLeuMetAsnThrGluArgThrAsnIleGlnValThrValThrGly 1211  
 Db 3875 GCTAAGTATGCCAGGCCACATTCATGACAAAGGAGATGTAAACATCAGTCAAGTCC 3934  
 QY 1212 ProSer-----SerPro 1215  
 Db 3935 AAGACTGATTCATCAACAGTTCATATGTGATCAAAATAAACGGGCTCTCTCTACAGAA 3994  
 QY 1216 SerProLeuAlaValAlaGlnProMetAlaValAsnIleSerAlaAsnGlyPheGlyPhe 1235  
 Db 3995 TCTCATCTGTCAATATC---CTGGGGAATCACTCTGTGTACGACGAGGAGGAGGCTGT 4051  
 QY 1236 AlaIleCysGlnLeuAsnValValTyrAsnValLysAlaSerGlySerSerArgArgArg 1255  
 Db 4052 GTGTATGTACAGAGTGTCTGTAGATACACATTCCTCCACT----- 4093  
 QY 1256 ArgSerIleGlnAsnGlnAlaPheAspLeuAspValAlaValLysGluAsnLys--- 1274

```

Db 4094 -----AGAACTGATGCCACCTTCTCACTAACAAGTGAAGACACTACGACAAATGT 4144
Qy 1275 -----AspAapLeuAenHleValAspLeuAenValCysThrSerPheSerGlyPro--- 1291
Db 4145 CCTCAAGATCTATTAAATTAATTAATGACCTTACGACTCTGTAATATACAGGAACCTCG 4204
Qy 1292 G1yAArgSerGlyMetAlaLeuMetG1yValAsnLeuSerGlyPheMetValProSer 1311
Db 4205 GAGAAATCAACATGCTGTGATTGAGTGAAGATGCTCTCTGATTT---ATCCCATCA 4261
Qy 1312 G1yAla-----IleSerLeuSerGluThrValLysLysValG1yTyrAspHis 1327
Db 4262 AAGGCTCAATCAGATCTGGAAGAGCAACCAATTAAGCCGAGTGAATTCAGACA 4321
Qy 1328 G1yLysLeuAenLeuTyrLeuAapSerValAsnG1yThrGlnPheCysValAsnIlePro 1347
Db 4322 GACATGTCATTAATTAATTAATGATGAGCTGGGACATGAGCTCTCCATCTGCTTCACG 4381
Qy 1348 AlAValArgAenPheLysValSerAsnThrGlnAspAlaSerValSerIleValAspTyr 1367
Db 4382 GTGACACAGACATTTGATGTAAGACCTTAAGCCGACAGTGAAGTATATAT 4441
Qy 1368 TyrGlnProArgGlnAlaValArgSerTyrAsnSerGluValLysLeuSerSerCys 1387
Db 4442 TATGAGACAGAGAGAGGACAGTTCGCCGAGTACAACTCC----- 4480
Qy 1388 AspLeuCysSerAapValGlnGlyCysArgProCys---G1yAapG1yAlaSerGly--- 1405
Db 4481 -----CCCTGCAGCTCTGATGGAGAGAAAGGAC 4510
Qy 1406 -----SerHis 1408
Db 4511 ACAAAATGAGATTCCTCCGCACTATTACAGAAACATAGTGAGCAAGTATCAAC 4570
Qy 1409 HisSerValIlePheIlePheCysPheLysLeuTyrPhe 1423
Db 4571 AGATCTCTGCAATATATTTTACATATATGAATTAATATATGTTT 4615

RESULT 4
BX391819/c 941 bp mRNA linear EST 13-MAY-2003
LOCUS BX391819 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
DEFINITION clone CS0D1081YB18 3-PRIME, mRNA sequence.
ACCESSION BX391819
VERSION BX391819.1 GI:30627404
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 941)
Full-length cDNA libraries and normalization
Unpublished
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 Evry cedex - France
Email: seqref@genoscope.cns.fr, web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 3578.f for
more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CS0BA1029ZD07 CS02740.1&cluster=3578.f.
Contact : Feng Liang Email : fliang@life.techn.com URL :
http://fulllength.invitrogen.com/Invitrogen Corporation 1600
Paradise Avenue Genoscope sequence ID : CS0BA1029ZD07_CS02740_1.
FEATURES
Source
1..941
location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0D1081YB18"

```

```

/tissue_type="PLACENTA COT 25-NORMALIZED"
/clone_id="Homo sapiens PLACENTA COT 25-NORMALIZED"
/notes="1st strand cDNA was primed with a NotI-oligo (dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."
BASE COUNT 271 a 214 c 181 g 266 t 9 others
ORIGIN
Alignment Scores:
Pred. No.: 6 01e-146 Length: 941
Score: 1461.00 Matches: 293
Percent Similarity: 93.95% Conservative: 2
Best Local Similarity: 93.31% Mismatches: 17
Query Match: 19.88% Indels: 3
DB: Gaps: 0
US-10-020-095-4 (1-1428) x BX391819 (1-941)
Qy 941 LysLysLysGlnLeuThrAapAenLeuLysG1yAlaLeuSerPheMetArgGlnGly 960
Db 938 AAAAGNANCANCTGACAGATATTTGAAAAGAAAGCTGTTTCATTATGAGGCAAGGG 879
Qy 961 TyrGlnArgGluLeuLeuTyrGlnArgGluAspGlySerPheSerAlaPheGlyAsnTyr 980
Db 878 TACCAGAGACACTTCTCTATCAGAGGAGATGGCTC-TTCACTGCTTTTGGAAATAT 820
Qy 981 AspProSerGlySerThrTrpLeuSerAlaPheValLeuArgCysPheLeuGluAlaAsp 1000
Db 819 GACCTCTGGAGCACTGTGTGTGCTGCTTGTGTTTAAGAGTTCCTTGAAAGCCGAT 760
Qy 1001 ProTyrIleAspIleAspGlnAenValLeuHisArgThrTyrThrTrpLeuLysGlyHis 1020
Db 759 CTTTACATATATATTTGATGGAATGTGTTACAGAAACATACACTTGCTTAAAGACAT 700
Qy 1021 GlnLysSerArgGlyGluPheTrpAapProGlyArgValIleHisSerGluLeuGlnGly 1040
Db 699 CAGAAATCCAAACGTTGATTTGGATCCAGAAAGTGATATATGATGAGCTTCAAGT 640
Qy 1041 G1yAsnLysSerProValThrLeuThrAlaTyrIleValThrSerLeuLeuGlyTyrArg 1060
Db 639 GGCATTAATAAGTCCAGTACACTTACGCTTAATATGATCTCTCTCGGATATAGA 580
Qy 1061 LysTyrGlnProArgAenLeuAapValGlnGluSerIleHisPheLeuGluSerGluPheSer 1080
Db 579 AAGTATCAGCTTAACATTTGATGCAAGAGTCAATTTTGGAGTCAATTCAGT 520
Qy 1081 ArgGlyIleSerAapAenTyrThrLeuAlaLeuIleThrTyrAlaLeuSerSerValGly 1100
Db 519 AGAGGAATTCACAGACATTTACTTACGCCCTTATTAATTATGATGTCATCAGTGGG 460
Qy 1101 SerProLysAlaLysGluAlaLeuAenMetLeuThrTyrArgAlaGlnGlnGlnGly 1120
Db 459 AGTCTTAAGCCGAGGAAGCTTTGAAATATGCTACTTGGAGACAGACAGAAAGGAGG 400
Qy 1121 MetGlnPheThrValSerSerGluSerLysLeuSerAapSerTrpGlnProArgSerLeu 1140
Db 399 ATGCAATTCGGGTGTCATGAGTCCAAACTTTCGACTCCCTGGAGCAGCCTCCCTG 340
Qy 1141 AspIleGluValAlaAlaTyrAlaLeuLeuSerHisPheLeuGlnPheGlnThrSerGlu 1160
Db 339 GATATGAGATTGACACCTTATGACAGTCTCTCACACTTCTTACAATTCAGACTTCGAG 280
Qy 1161 G1yTyrLeuProIleMetArgTrpLeuSerArgGlnArgAsnSerLeuGlyIlePheAlaSer 1180
Db 279 GGAATCCCAATTTATGAGTGGCTTAACAGCAAGAAATAGCTTGAGGTGTTTGCACTCT 220
Qy 1181 ThrGlnAspThrThrValAlaLeuLysAlaLeuSerGluPheAlaAlaLeuMetAenThr 1200
Db 219 ACTCAGATTCACCTGTGCTTTAAAGGCTCTGTGCAATTCAGCCCTTAAGATACA 160
Qy 1201 G1yArgThrAsnIleGlnValThrValThrGlyProSerSerProSerProLeuAlaVal 1220

```



Db 159 GAAAGACAATAATCCAGTACCGTGACGGGCGCTAGCTCACCAAGTCTCTTGCTGTG 100  
 QY 1221 ValGlnProMetAlaValAsnIleSerAlaAsnGlyPheGlyPheAlaIleCysGlnLeu 1240  
 Db 99 GTACAGCCACAGCGCACTTAATATTCGCCAAATGTTGGATTGGATTGCTATTGTCAGCTC 40  
 QY 1241 Asn--ValValIlyrAsnValIlyrAsnGlySerSer 1252  
 Db 39 AATGNTTGTATATATGTAAGCGCTTGCTGCTCTCT 2  
 RESULT 5  
 AL554092 1201 bp mRNA linear EST 31-MAY-2003  
 LOCUS AL554092 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA  
 DEFINITION clone CS0D1081YB18 5-PRIME, mRNA sequence.  
 AL554092  
 ACCESSION AL554092.2 GI:31275905  
 VERSION  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 REFERENCE 1 (bases 1 to 1201)  
 AUTHORS Li, W.B., Gruber, C., Jesse, J. and Polayes, D.  
 TITLE Full-length cDNA libraries and normalization  
 JOURNAL Unpublished  
 COMMENT On Feb 15, 2001 this sequence version replaced gi:12894544.  
 CONTACT Genoscope  
 Genoscope - Centre National de Sequencage  
 BP 191 91006 Evry cedex - France  
 Email: seqref@genoscope.cns.fr; Web : www.genoscope.cns.fr  
 Library was constructed by Life Technologies, a division of  
 Invitrogen. This sequence belongs to sequence cluster 3578.f For  
 more information about this cluster, see  
 http://www.genoscope.cns.fr/  
 cgi-bin/cluster.cgi?seq=CS0D1081YB18&ql=3578.f. Contact :  
 Feng Liang Email : fliang@life.techn.com URL :  
 http://fulllength.invitrogen.com/ Invitrogen Corporation 1600  
 Faraday Avenue Genoscope sequence ID : CS0D1081YB18QPL.  
 FEATURES  
 source  
 1..1201  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="CS0D1081YB18"  
 /tissue\_type="PLACENTA COT 25-NORMALIZED"  
 /clone\_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"  
 /note="1st strand cDNA was primed with a NotI-oligo (dT)  
 primer. Five prime end enriched, double-strand cDNA was  
 digested with Not I and cloned into the Not I and EcoR V  
 sites of the pCMVSPORT 6 vector. Library was normalized."  
 BASE COUNT 359 a 238 c 236 g 293 t 75 others  
 ORIGIN  
 Alignment Scores:  
 Pred. No.: 7.4e-140 Length: 1201  
 Score: 1406.00 Matches: 293  
 Percent Similarity: 90.74% Conservative: 1  
 Best Local Similarity: 90.43% Mismatches: 29  
 Query Match: 19.13% Indels: 4  
 DB: 9 Gaps: 1  
 US-10-020-095-4 (1-1428) x AL554092 (1-1201)  
 QY 1 MetGlnGlyProCProLeuLeuThrAlaAlaHisIleLeuLeuCysValCysThrAlaAlaLeu 20  
 Db 92 ATGCGAGGGGCCACCGCTCTGAMCGCMGCCACCTCTCGTGTCGACCGCCGCGCTG 151  
 QY 21 AlaValAlaProGlyProAlaPheLeuValThrAlaProGlyIleIleLeuProGlyGly 40  
 Db 152 GCGCGAGGCTCCCGGGGCTGCTTTCGGGAMAGCMCCAGGATATATCAAGCCCGAGGCA 211  
 QY 41 AsnValThrIleGlyValGlnLeuLeuGlnHisCysProSerGlnValThrValIlyAsn 60

Db 212 AATGACTATATGGGGGTGGAGCTTCTGGAAATACGCGCTTAMAGTACTGTGAAGCG 271  
 QY 61 GlnLeuLeuysThrAlaSerAsnLeuThrValSerValIleGlnAlaGlnGlyAlaPhe 80  
 Db 272 GAGCTGCTCAAGACGACGATCAAAAMTCACTGCTCTGTCTGAGGAGGAAAGGAGCTCTT 331  
 QY 81 GlnIlyGlySerPheIlyeThrIleuThrleuProSerleuProLeuAsnSerAlaAspGlu 100  
 Db 332 GAAAAAGCTCTTTTAAAGAACTTATCTTCCATTAATACMTCTGAACACATGCAATGAG 391  
 QY 101 IleTyGlnLeuArgValThrGlyArgThrGlnAspGlnIleLeuPheSerAsnSerThr 120  
 Db 392 ATTATGAGCTAGMTAGTGAACCGAGCTCCAGATGAGATTTATCTCTTAAGTAC 451  
 QY 121 ArgLeuSerPheGlnIlyThrIlyArgIleSerValPheIleGlnThrAspIlyAlaLeuTyr 140  
 Db 452 CGCTTATCATTTGAGACCAAGAGAAATATCTGTTTCATTCAAAACAGAAAGGCTTATAC 511  
 QY 141 LysProIlySGlnGlnValIlyPheArgIleValThrIleuPheSerAspPheIlyProTyr 160  
 Db 512 AAGCAAAAGCAAGAGTAAAGTTTGCATTGTACACTCTTCTCAGATTTTAAAGCTTAC 571  
 QY 161 LysThrSerIleuAsnIleLeuIleIlyAspProIlySerAsnLeuIleGlnGlnThrPhe 180  
 Db 572 AAAACCTCTTAAACATTCCTCATTAAGACCCCAATCAATTGATCAACAGGTG 631  
 QY 181 SerGlnIlySerAspIlyValIleSerIlyThrPheGlnLeuSerSerHisProIle 200  
 Db 632 TCACAAACAAATGTAGCTTGAGTCAATTCACAAACTTTCAGCTATCTTCCATCCAAATA 691  
 QY 201 LeuGlyAspIlyPheIleGlnValGlnValAsnAspGlnThrTyThrGlnSerPheGln 220  
 Db 692 CTGGTGTAGCTGTCTATTCAAGTTCAGTGAATGACGACGACATTAATCAATCATTTAG 751  
 QY 221 ValSerGlnTyThrValLeuProIlyPheGlnValThrLeuGlnThrProLeuTyCysSer 240  
 Db 752 GTTTCAGATATATGATATTCACAAATTTGAAGTCACTTTCAGACACCATTAATATTGTTCT 811  
 QY 241 MetAsnSerIlyHisIleAsnGlyThrIleThrAlaIlyTyThrTyGlyIlyAspProVal 260  
 Db 812 ATGAATTCCTAAGCATTTTAATGTATATC--ACGGCAAGTATATCATATGGAACGAGTG 870  
 QY 261 LysGlyAspValThrIleuThrPheLeuProIlySerPheTyGlyIlyAsnIle 280  
 Db 871 AAAGAGACGTAAGCTTCAATTTTACCTTTTCTTTGAGAAAGAAATAATAT 930  
 QY 281 ThrIlyThrPheIlyIleAsnGlySerAlaAsnPheSerPheAsnAspGlnIlyMetIlyS 300  
 Db 931 ACAAAACATTTAAGATTAATGAATGATCTGCACAACTTCTTTTATGAGAAAGATGAAA 990  
 QY 301 AsnValMetAspSerSerAsnGlyLeuSerGlnTyThrIleuAspLeuSerSerProGlyPro 320  
 Db 991 AM-GTWTATGATTC-TCAATATGACTTCTGAA--ACYKAGATATATTCCTCCGACM 1045  
 QY 321 ValGlnIleLeu 324  
 Db 1046 GTAGAAATTTTA 1057  
 RESULT 6  
 BC049868 4667 bp mRNA linear HTC 01-APR-2003  
 LOCUS BC049868  
 DEFINITION Mus musculus, murinoglobulin 1, clone IMAGE:5097177, mRNA.  
 ACCESSION BC049868  
 VERSION BC049868.1 GI:29437204  
 KEYWORDS HTC.  
 SOURCE Mus musculus (house mouse)  
 ORGANISM Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 REFERENCE 1 (bases 1 to 4667)  
 AUTHORS Strausberg, R.  
 TITLE Direct Submission

## JOURNAL

Submitted (31-MAR-2003) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA

REMARK  
COMMENT  
NIH-MGC Project URL: <http://mgc.nci.nih.gov>  
Contact: MGC help desk  
Email: [cgabs-remail.nih.gov](mailto:cgabs-remail.nih.gov)  
Tissue Procurement: Jeffrey E. Green, M.D.  
cDNA Library Preparation: Life Technologies, Inc.  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LIML)  
DNA Sequencing by: National Institutes of Health Intramural Sequencing Center (NISC),  
Gaithersburg, Maryland;  
Web site: <http://www.nisc.nih.gov/>  
Contact: [nisc.mgc@nih.gov](mailto:nisc.mgc@nih.gov)  
Ahrens, N., Ayle, K., Beckstrom-Sternberg, S.M., Benjamin, B., Blakesley, R.W., Bouffard, G.G., Breen, K., Brinkley, C., Brooks, S., Dietrich, N.L., Granite, S., Guan, X., Gupta, J., Haghighi, P., Hansen, N., Ho, S.-L., Karlins, E., Kwong, P., Latic, P., Legaspi, R., Madwed, Q.L., Mastello, C., Maskeri, B., Mastrian, S.D., McCloskey, J.C., McDowell, J., Pearson, R., Stantirpop, S., Thomas, P.J., Touchman, J.W., Turgerson, C., Vogt, J.L., Walker, M.A., Wetherby, K.D., Wiggins, L., Young, A., Zhang, L.-H. and Green, E.D.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LIML at: <http://image.liml.gov>  
Series: IRAC Plate: 99 Row: h Column: 5  
This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 6678965  
This clone has the following problem: frame shifted.

## FEATURES

## source

1. 4667  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="FVB/N"  
/db\_xref="taxon:10090"  
/clone="IMAGE:5097177"  
/tissue\_type="Liver, normal, 5 month old male mouse."  
/clone\_id="NCI CGAP L4.9"  
/lab\_host="DH10B"  
/note="Vector: pCMV-SPORT6"  
BASE COUNT 1346 a 1126 c 1046 g 1149 t  
ORIGIN

## Alignment Scores:

Pred. No.: 7 39e-139 Length: 4667  
Score: 1405.50 Matches: 445  
Percent Similarity: 45.10% Conservative: 268  
Best Local Similarity: 28.15% Mismatch: 559  
Query Match: 19.13% Indels: 310  
DB: 11 Gaps: 53

US-10-020-095-4 (1-1428) x BC049868 (1-4667)

QY 4 ProProleuLeuThraAlaAlaLeu-----LeuCyValCys 16  
DB 73 CTTCTGCTCTTCTTACACACATGTGGAAGAGAGAGGGCTGAGTGTGCTTTT 132  
QY 17 ThraAlaAlaLeuA-----ValAlaProGlyProAlaGlyPheLeu 29  
DB 133 TCGGTTCTCTCTCTGCTTCGATCTGCTTTTACTCAACGAGACTGGAATATATG 192  
QY 30 ValThraAlaProGlyLeileargProGlyGlyAsnValThrlleGlyValGluLeu 49  
DB 193 GTGCTGTCTCTTCT-----CACTTAC 216  
QY 50 GluHisCysProSerGlnValThraValGluLeuLeuGlyThraAlaSerAsnLeu 69  
DB 217 ACTGAGACCCCTGAGAAATCTGCTCATCTATACAGCTGAATGAACAGTACCGTC 276  
QY 70 ThraValSerValLeuGluAlaGluGly-----ValPheGlu 81  
DB 277 ACAGCTTCTTGTATCTCAATCAGAAAGAAACCTGTTGATGAGTGTGCTTAC 336

QY 82 LysGlySerPheLysThrlLeuThr-----LeuProSerLeuProLeuAsnSerAlaAsp 99  
DB 337 AAGGACTGTGTTCACAAAGTGTCTCTTCAATATCCCTAGACTCTCTCTCTGTATGAGAG 396  
QY 100 GluIleTyrGluLeuArgValThrglyArgThrglnAspGluIleLeuPheSerAsnSer 119  
DB 397 GATTTCTCTTACTGACATCAAAAGGCCAACCCAGAA-----TTCAGCAAAAGG 447  
QY 120 ThraGluSerPheGluThrlLysArgIleSerValPheIleGlnThraAspValAlaLeu 139  
DB 448 AAGCAGTCTGTGTGAAGAACAAAGAAAGTGTGTCTTGTGTGACAGATTAAGCCCGTC 507  
QY 140 TyrLysProLysGlnValLysPheArgIleValThrlLeuPheSerAspPheLysPro 159  
DB 508 TACAAACCCAGACAAATCAGTTAAATTCGGGTGTCTATGATTAATATGACTACCTCC 567  
QY 160 TyrLysThrlSerLeuAsnIle-----LeuIleLysAspProLysSerAsnLeuIleGln 178  
DB 568 CTGAATGAGTTGCTTCTGCTTACATTTGAGGATCCAAAGAAAGCAATATATGACAG 627  
QY 179 TrpLeuSerGlnIleSerAspLeuGlyValIleSerLysThrlPheGlnLeuSerSerHis 198  
DB 628 TGGAGGATTTTAAGACTGAGAAATGGGCTTAAGCAATGCTTCAAGCTTGGACAGAG 687  
QY 199 ProIleLeuGlyAspTrpSerIleGlnVal-----GlnValAsnAspGlnThrlTyrGln 217  
DB 688 CCATTTCAGGGCCCTTACAAAGATATGATGTGTGACAAAGATGACGGGAGAAAGAAAC 747  
QY 218 SerPheGlnValSerGluTyrValLeuProLysPheGluValThrlLeuGlnThrlProLeu 237  
DB 748 TCTTTTACTGTGATGAAATTTGTGCTTCCAGATTTTAATGTCGACTGAGATCCCAAT 807  
QY 238 TyrCysSerMetAsnSerLysHisLeuAsnGlyThrlleThraAlaLysTyrThrlGly 257  
DB 808 GCCATGCTGTGATATGATGAGTCTCAGTGTGACTGATGTGGAAATATACCTATGGG 867  
QY 258 LysProValLysGlyAspValThrlLeuThrlPheLeuProLeuSerPheTrpGlyLys 277  
DB 868 AAGCTGTTCAGACATGTG----- 888  
QY 278 LysAsnIleThrlLysThrlPheLysIleAsnGlySerAlaAsnPheSerPheAsnAspGlu 297  
DB 889 -----AAGATTAACGTCTGCGGTGAGACTGAGACTGAGTACAG 927  
QY 298 GluMetLysAsnValMetAspSerSerAsnGlyLeuSer-----GluTyrLeu 313  
DB 928 GAAGTAAATTTCTAGCTAGAC--AACAATGGCTGACATACACAAAGTGAATCACT 984  
QY 314 AspLeuSerSerProGlyPro-----ValGluIleLeuThrlVal 327  
DB 985 GAGCTCAATCAAAAGAAAGAAATATGAGTTACAGCTTTTCCATGTGAATGCACTGTT 1044  
QY 328 ThrlGluSerValThrlGlyLe-----SerArgAsnValSerThraAsnValPhePheLys 345  
DB 1045 ACAGAGAAAGGAGCAGAGATTCAGTTCAGTTCGTAACAATTAATTAAGAAATC 1104  
QY 346 GlnHisAspTyrIleIleGluPhePheAspTyrThrlValLeuLysProSerLeuAsn 365  
DB 1105 ACAAAAGACTCTA-----TTCTGAAGGACAGATTCACACTTCAAGACATGGAATTC 1158  
QY 366 PheThrlAlaThraValLysValThraArgAlaAspGlyAsnGlnLeuThrlLeuGluArg 385  
DB 1159 TTTCTTGTGAAGTCCGCTTATGATATCAAGAGACCCATATCCAAATAGAGAA----- 1215  
QY 386 ArgAsnAsnValIleThraValThrglnArgAsnTyrThrlGluTyrTrpSerGlySer 405  
DB 1216 -----GTCTTATCAAGACAAAGAACTAGCTATACAGTCTACTACACATGAT 1266  
QY 406 AsnSerGlyAsnGlnLysMetGluAlaValGlnLysIleAsnTyrThrlValProGlnSer 425  
DB 1267 CAGCATGGC-----CTG 1278

QY 426 GlyThrPheIysIleGluPheProIleLeuGluAspSerSerGluLeuIleuLysAla 445  
Db 1279 GCAGAGTCTCCATGATGATACCAAGCATCTCGGGCTCTTCC---CTCATATCAAAGTC 1335  
QY 446 -----TyrPheLeuGlySerLysSerMet 454  
Db 1336 AACCCAAAGAGGAAGATTCAATGTTCTATTCTATGTCATGAGAGAACATCCAGT 1395  
QY 455 AlaValHisSerLeuPheLys-----SerProSerLysThrTyrIleGluLeuLysThr 472  
Db 1396 GCAAGACATGTCGCTATGCTGTTTACTCTCCACAGAGAGCTATCACTTACTGACACA 1455  
QY 473 ArgAspGluAsnIle-----LysValGlySerProPheGluLeu 485  
Db 1456 GAGACAGACGATCTTGCCTGCAACACGATTCAACAGTTCAGGACATTTT----- 1509  
QY 486 ValValSerGlyAsnLysArg---LeuLysGluLeu-----SerTyrMetAlaValSer 502  
Db 1510 ATTCTGAAGGGGAGCTTGGAGTGTGGAAGAGCTCATTTTCTACTACCTGTGCATGGCG 1569  
QY 503 ArgGlyGlnLeuValAlaValGlyLysGlnAsnSerThrMet----- 516  
Db 1570 CAGGAGACATCATCCAGACTGGAACCATCTACACAGTGTGAGCCAGAGAACCTCCA 1629  
QY 517 -----PheSerLeuThrProGluAsnSerTyrThrProLysAlaCys 530  
Db 1630 GTAAAGAGAAATTTGCTTGAGATCCCTGTGAGTTTATGATGATGCTCCCATGCTAA 1689  
QY 531 ValIleValTyrTyrIleGluAspAspGlyGluIleIleSerAspValLeuLysIlePro 550  
Db 1690 AGTCATCTTACACATCTTGCCTATGAGAGATGAGATTCGTGTAACCTTTGAA 1749  
QY 551 ValGlnLeuValPheLysAsnLysIleLysLeuTyrTyrSerLysValIleGluPro 570  
Db 1750 ATTGAAGAGTCTTCTGCAACAAGTGAAGCTGCGCTTACAGCATTTCAAGTCTTCC 1809  
QY 571 SerGluLysValSerLeuArgIleSerValThrGlnProAspSerIleValGlyIleVal 590  
Db 1810 GCCTCAACAACCCGCTGCAAGAGTCAAGCTTCT---CTCAGTCCCTGTGTGCGCTGAGA 1866  
QY 591 AlaValAspLysSerValAsnLeuMetAsnAlaSerAsnAspIleThrMetGluAsnVal 610  
Db 1867 GCTGTGACCAAGAGTGTCTACTCTGTAACCCGAGTGTGAGCTCTCCCTTCC----- 1920  
QY 611 ValHisGluLeuGluLeuTyrAsnThrGlyTyrTyrIleuGlyMetPheMetAsnSerPhe 630  
Db 1921 -----TGAATATACATCTGCCA-----GGTATGACGACAAACAAATTC 1959  
QY 631 AlaValPhe-----GlnGluCysGlyLeu-----TyrVal 640  
Db 1960 GTTCCAGTCCCGTCTGTCTGAAGACCAAGAACTGTATCTGTACAGTTCACTGTTG 2019  
QY 641 -----Leu 641  
Db 2020 GCTGAGAACACAACTTAGTACACATGGAACGTGAGAGAGATGTCTATGATATG 2079  
QY 642 ThrAspAlaAsnLeuThrLys----- 648  
Db 2080 GAGGACATGGGTTTAAACGATTCACCAACTGATGATCAAACTTCTATAATTTGTTT 2139  
QY 649 AspTyr----- 650  
Db 2140 GACTATGGAGTGTCCAAATATCAGACCTCGTGTGAATTTGATCTTGACCTTACCCCT 2199  
QY 651 ---IleAsp-GlyValTyrAspAsnAlaGluTyrAlaGluLysPheMetGluLysAsnG 669  
Db 2200 GAAATTTCTTGGAGTT-----TACGACACACATTAGTAAAGACTGAA 2244  
QY 669 uclYHisIleValAspIleHisAspPheSerLeuGlySerSerProHisValArgLysHis 689  
Db 2245 GAGCCACCAAGTAAAGATTCATCATGAAATGATCCA-CTTACAGAGACCATTCGAAANA 2303  
QY 689 sPheProGluThrTyrIleTyrLeuAspThrAsnMetGlyTyrArgIleTyrGlnLys 709

Db 2304 TTTTCTGTAACCTGGGCTGGGATATAGTACAGTAACTCCACAGACTGCTGAAGT 2363  
QY 709 eGluValThrValProAspSerIleThrSerTyrValAlaThrGlyPheValIleSerG 729  
Db 2364 GGAATGACTCTCTTACACATCATCTGAATGGAAGCAGAGCTCTCTGTTCCAA 2423  
QY 729 uAspLeuGlyLeuGlyLeuThrThrThrProValGluLeuGlnAlaPheGlnProPheP 749  
Db 2424 TGACATGCTGCTTGCGCTCTTCTGAG---GTACCTTCCAAAGCTTCAAGCCTTCTT 2480  
QY 749 eIlePheLeuAsnLeuProLysSerValIleArgGlyGluGluPheAlaLeuGluIleTh 769  
Db 2481 TGTGAGGCTCATTTGCTTATCTGTGTGCTGTGAGAAAGCTTCAAGCTCAAGGCCAC 2540  
QY 769 rIlePheAsnTyrLeuLysAspAlaThrGlnValLysValIleIleGlyLysSerAsp 789  
Db 2541 TGTGATGAATATCTTCCCAACAGCATGATGATGATGATGATGATGATGATGATGAT 2600  
QY 789 sPheAspIleLeuMetThrSerSerGluIleAsnAlaThrGlyHisGlnGlnThrLeu 809  
Db 2601 TTTCACTGCTCCAGTGGAGAGAC-----CAAGATCTTACTG 2642  
QY 809 uValProSerGluAspGlyAlaThrValLeuPheProIleArgProThrHisLeuGly 829  
Db 2643 CCTCAGTCCCAATGGAGGACACACTCATCTGCTGTGATCTCCAAAGCTTTAAGGAA 2702  
QY 829 uIleProIleThrValThrAla-----LeuSerProThrAlaSerAspAl 844  
Db 2703 TGTGAATTTCTGTGTGTGTGGAAGACAAACAGTCTCAGAGCTGTGTGTGTGAGT 2762  
QY 844 aValThr-----Gln-MetIleLeuValLysAlaG 854  
Db 2763 GCCCAGCTTCTGTAACCTGGAGAAAGACAGTAAAGTCAATGATTTAGACTG 2822  
QY 854 IugIyIleGluLysSerTyrSerGlnSerIleLeuLeuAspLeuThrAspAsnArgLeu 874  
Db 2823 AAGGATCAACAGACATGATCATCTGAGTCTGATGATGATGATGATGATGATGAT 2882  
QY 874 InSerThrLeuLysThrLeuSerPheSerPheProProAsnThrValThrGlySerGlu 894  
Db 2883 CT-----GAAAAATGCTCCTGTCTCCCAACAGTGTGTAAGATCAGCA 2933  
QY 894 rGValGlnIleThrAlaIleGlyAspValLeuGlyProSerIleAsnGlyLeuLase 914  
Db 2934 GAGCCATTTCTGTGTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2993  
QY 914 euIleArgMetProTyrGlyCysGlyGluGlnAsnMetIleAsnPheAlaProAsnIle 934  
Db 2994 TTCTCATATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3053  
QY 934 rYrIleLeuAspTyrLeuThrLysLysGlyGlnLeuThrAspAsnLeuLysGlyLysAla 954  
Db 3054 ACGTACTGAATATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3113  
QY 954 euSerPheMetArgGlnGlyTyrGlnArgGlyLysLeuTyrGlnArgGlyLysSerP 974  
Db 3114 TTGGTCTCTGAGAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3173  
QY 974 heSerAlaPheGly-----AsnTyrAspProSerGlySerThrTyrLeuSerAlaPhe 992  
Db 3174 ACAAGTCTTGGGGATGAAATGGCGAAGAGAGAGAAACATCTGCTCAACGCTTTG 3233  
QY 992 alleuArgCysPheLeuGluAlaAspProTyrIleAspIleAspGlnAsnValLeuHis 1012  
Db 3234 TGTCTCAAGTCTTGGCCCAAGCTGAGCTTCACTTATGATGATGATGATGATGATGAT 3293  
QY 1012 rGlnThrTyrThrTyrLeuLysGlyHisGlnLysSerAsnGlyGluPheThrAspProG 1032  
Db 3294 ATGCTTCACTGCTGCTCCCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3353  
QY 1032 rGValIleHisSerGluLeuGlnGlyLeuLysSerProValThrLeuThrAlaTyr 1052

```

Db 3354 CATTGTTCAACAATGCCATGAGGGGAGTAGATGATGAATGAAATGACCTCTCTGCTTACCA 3413
Qy 1052 levalThrSerLeuLeuGlyTyrrArgLysTyrrGlnProAsnIleAspValGlnIleSer 1072
Db 3414 TAACCATGCGCTTCTTGAAAAGTTTCACTCCAGCCAGCCATCCGTGTTGCTCCAAAGCCC 3473
Qy 1072 leHisPheLeuGluSer-----GluPheSerArgGlyIleSerAspAsnT 1087
Db 3474 TGACCTGCTGAGCATCTCTGAAGACCATATAGAACAGAAAGCAAGAAATGCCAGCTTGTGT 3533
Qy 1087 TyrrThrLeuAlaLeuIleThrTyrrAlaLeuSerSerValGlySerPro---LysAlaIle 1106
Db 3534 ACACCAAGCGCTTATGAGCTTATGCTTCTGCTGCGCGGAGAACCAAGATTAAGAGAGATG 3593
Qy 1106 lualAlaLeuAsnMetLeuThrTrpArgIleGlnGlnGlyGlyMetGlnPheTrpVal- 1125
Db 3594 AAATCTGTAAGTCACTTATGATGAGAGAGTATTAAGAAAGAAACAATCTATCCATCGGAAA 3653
Qy 1126 --SerSerGluSerLysLeuSerAspSerTrp-----GlnProArgSerLeuA 1141
Db 3654 GACCTCAAGAAATCCAGAAATCTGAAACACCATTTGTACAAACCCAGGCTTCTCTGCTG 3713
Qy 1141 spIleGlnValAlaAlaThrTyrrAlaLeuLeuSerHisPheLeuGlnPheGlnThrSerGlu 1161
Db 3714 AAGTAAAGATGAATGATATATGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3761
Qy 1161 LysIlePro-----IleMetArgTrpLeuSerArg 1171
Db 3762 CAGCCCACTCCCTGAGAGATCTGACTTTGTCAATGAGACATCATGTGCTCACAAGC 3821
Qy 1171 lnaArgAsnSerLeuGlyGlyPheAlaSerThrGlnAspThrTrpValAlaLeuLysAla 1191
Db 3822 AACGAAATTCGAATGTGGTCTTCTCTCCACACAGACAGACTGTGTGGCCCTTGATGCTC 3881
Qy 1191 eusSerGlnPheAlaAlaLeu-----MetAsnThrGlnArgThrAsnIleGlnValT 1208
Db 3882 TGTCCAATATGAGCAGAGTACTTTTTCGAGAAAGTCAAGAAATCTTGTGTGATCTTCC 3941
Qy 1208 hrValThrGlyProSerSerPro-----SerProLeuAlaValValG 1222
Db 3942 AATCTACAGGGTCACTTTTCCAAAAGTTCCAAAGTGAAGAACAGTAATCGCTGTACTAC 4001
Qy 1222 lnaProMetAlaVal-----AsnIleSerAlaAsnGlyPheGly 1235
Db 4002 AGCAGGTGCGATTCACAGACATTCCTGGGACATACATCATGATGTGACAGGAGAGAT 4061
Qy 1235 heaAlaIleCysGlnLeuAsnValValTyrrAsnValLysAlaSerGlySerSerArgArg 1255
Db 4062 GTGTGTATGCTCAGATATGCTGAGATACAAACATGCACTTGAG----- 4105
Qy 1255 rgaArgSerIleGlnAsnGlnGlnAlaPheAspLeuAspValAlaValLys----- 1271
Db 4106 -----AACGAGCTGTCTGCAATTTGCTATATGGGTCACAGACAGTACTTAACCTT 4154
Qy 1272 --GluAsnLysAspAspLeuAsnHisValAspLeuAsnValCysThrSerPheSerGlyP 1291
Db 4155 GTAATTAACCCCAAGGCGACAAACAGCTTCAGATCTCACTAGAAATCAGATTACAGACGA 4214
Qy 1291 ro---GlyArgSerGlyMetAlaLeuMetGlnValAsnLeuLeuSerGlyPheMetValP 1310
Db 4215 GCCCGCAGAGCTTCAACATGTGTATGCTGATGTGAAGATGCTATCTGGTTTC----- 4267
Qy 1310 roserGlnAlaIleSerLeuSerGlnThrValLysLysVal----- 1323
Db 4268 -----ATCCCATTTGAACCAACAGTGAAGAAAGCTTGAAAGATTAGACACAGCTGA 4316
Qy 1324 -----GluTyrrAspHisGlyLysLeuAsnLeuTyrrLeuAspSerValAsnGlnThr 1341
Db 4317 GCAGAACAGAGTAGAGCAACAATGCTTAATATATTTTGCATCAGGTGACCAATGACA 4376
Qy 1341 lnaPheCysValAsnIleProAlaValArgAsnPheLysValSerAsnThrGlnAspAla 1361
Db 4377 CACTGGCTTCTCTTCAATCATTCACAAAGATATCCAGTAAAGAACTGTGAGAGCTGCA 4436

```

```

Qy 1361 exValSerIleValAspTyrrGlyProArgGlnAlaValArgSerTyrrAsnSer 1380
Db 4437 TTGTGAAGATCTATGATCTATGAGACAGATGAATGAGCTTTTGTGATACAGACAGC 4495
RESULT 7
BUS17328
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

```

```

BUS17328 879 bp mRNA linear EST 12-SEP-2002
AGENCOURT 10164521 NIH_MGC_71 Homo sapiens cDNA clone IMAGE:6514905
5', mRNA sequence.
BUS17328
BUS17328.1 GI:22824854
EST.
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
1 (bases 1 to 879)
NIH-MGC http://mgi.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cga@bbs-r@mail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLNL4090 row: P column: 10
High quality sequence stop: 663.
Location/Qualifiers
1..879
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6514905"
/tissue_type="leiomyosarcoma"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_71"
/note="Organ: uterus; Vector: pCMV-SPORT6; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 2.1 kb."

```

```

BASE COUNT 248 a 171 c 206 g 253 t 1 others
ORIGIN
Alignment Scores:
Pred. No.: 1,64e-126 length: 879
Score: 1280.00 Matches: 263
Percent Similarity: 91.38% Conservative: 2
Best Local Similarity: 90.69% Mismatches: 6
Query Match: 17,428 Indels: 19
DB: 13 Gaps: 2
US-10-020-095-4 (1-1428) x BUS17328 (1-879)

```

```

Qy 1070 GluSerIleHisPheLeuGluSerGluPheSerArgGlyIleSerAspAsnTyrrThrLeu 1089
Db 2 GACTCTATCACTATTTTGGAGTCTGAATTCAGTGAAGAAATTCACAAATATATCTTA 61
Qy 1090 AlaLeuIleThrTyrrAlaLeuSerSerValGlySerProLysAlaIleGlnAlaLeuAsn 1109
Db 62 GCCCTTAATACCTTATGATCTATGATCAGTGGGAGTCTTAACCGAAGAGCTTGAAT 121
Qy 1110 MetLeuThrTrpArgIleGlnGlnGlnGlyGlyMetGlnPheTrpValSerSerGluSer 1129
Db 122 ATGCTGATTTGAGAGACAAACAGAGGTGGCATCATTTCTGGGTGTCATCAGAGTCC 181
Qy 1130 LysLeuSerAspSerTrpGlnProArgSerLeuAspIleGlnValAlaAlaThrAlaLeu 1149
Db 182 AAATCTTGAATCTCCGAGACCAAGCTCTCGAGATTAATGAAGTTGACACCTATGACATG 241

```

FEATURES	LOCATION/Qualifiers
Db	1150 LeuSerHisPheLeuGlnPheGlnThrSerGluGlyIleProIleMetArgTrpLeuSer 1165
Db	242 CTCACACTTCTTCAATTTCACACTTTCAGCTTCAGGAAATCCCAATTATGAGTGCGTAAAGC 301
Qy	1170 ArgGlnArgAsnSerLeuGlyGlyPheAlaSerThrGlnAspThr-ThrValAlaLeuLys 1185
Db	302 AGCAAAAGAAATAGCTGGTGGTGTTTGGATCTACTACAGGATACCACTGTGGCTTTTAAAG 361
Qy	1190 AlaLeuSerGluPheAlaAlaLeuMetAsnThrGluArgThrAsnIleGlnValThrVal 1205
Db	362 GCTCTGCTGTAATTTGCAGACCTCTAATGAATACAGAAAGACAAATATCCAGTGCACGG 421
Qy	1210 ThrGlyProSerSerProSerPro----- 1217
Db	422 ACGGGGCTTAGCTACCAAGTCTCTTAAAGTTTCTGATTGACACACACACCGCTTACTC 481
Qy	1218-----LeuAlaValValGlnProMetAlaValAsnIleSerAlaAsnGly 1233
Db	482 CTTGACAGACAGACAGACTTGTCTGTGTGTACAGCCAAAGCGAGTTAATTTCCGCAATGCT 541
Qy	1233 PheGlyPheAlaIleCysGlnLeuAsnValValIlyAsnValLysAlaSerGlySerSer 1253
Db	542 TTTGATTTGCTAATTTGTTCAGCTCAATGTTGTAATGTAATGTAAGGCTTTCGGGCTTCT 601
Qy	1253 ArgArgArgArgSerIleGlnAsnGlnGlnAlaPheAspLeuAspValAlaValLysGlu 1273
Db	602 AGAAGACCAAGATCTATCCAAATACAGAAAGCCTTTGATTAAGATGTTGCTGTAAAGAA 661
Qy	1273 AsnLysAspAspLeuAsnHisValAspLeuAsnValCysThrSerPheSerGlyProGly 1292
Db	662 AATAAAGATGATCTCAATCACTGTGATTTGAATGTGTGTACAGCTTTTCGGGCCCGGCT 721
Qy	1293 ArgSerGlyMetAlaLeuMetGluValAsnLeuLeuSerGlyPheMetValProSerGlu 1312
Db	722 AGGAGTGCGACATGCGCTTATGGAAGTAACTAATTAAGGGGCTTATGCGCTTACAGA 781
Qy	1313 AlaIleSer-LeuSerGluThrValLysLysValGluIlyrAspHisGlyIlyrLeuAsnLe 1333
Db	782 GCAATTTCTCTGACCGAGACGAGTGAAGAAAGTGGGATATGATCATGGGAAAACTCAAC 841
Qy	1332 uTyrlau---AspSerValaasnGluThr 1340
Db	842 CTCATTTTACGATTTCTGTAAATGAAC 869
RESULT 8	
BQ028040/c	667 bp mRNA linear EST 27-MAR-2002
LOCUS	UI-H-C00-arh-c-08-0-UI.s1 NCI_GAP_Sub9 Homo sapiens cDNA clone
DEFINITION	IMAGE3106526 3', mRNA sequence.
ACCESSION	BQ028040
VERSION	BQ028040.1 GI:19763319
KEYWORDS	EST.
SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrate; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo. 1 (bases 1 to 667)
AUTHORS	NCI-GAP http://www.ncbi.nlm.nih.gov/ncigap.
TITLE	National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
JOURNAL	Unpublished
COMMENT	Contact: Robert Strausberg, Ph.D. Email: cgaabg-remail.nih.gov Tissue Procurement: Dr. Jose Mercuende cDNA library preparation: Dr. M. Bento Soares, University of Iowa cDNA library Arrayed by: Dr. M. Bento Soares, University of Iowa DNA Sequencing by: Dr. M. Bento Soares, University of Iowa Clone Distribution: Clone distribution information can be found through the I.M.A.G.E. Consortium/ILNI, at: http://image.llnl.gov Seq primer: M13 FORWARD POLYA+yes.

		/organism="Homo sapiens"						
		/mol_type="mRNA"						
		/db_xref="taxon:9606"						
		/clone="IMAGE:3106526"						
		/issue_type="mixed"						
		/dev_stage="mixed"						
		/lab_host="DH10B (Life Technologies)"						
		/clone_lib="NCI_CGAP_Sub9"						
		/note="Vector: pT73-Pac (Pharmacia) with a modified polylinker; Site 1: EcoR I; Site 2: Not I; tissues: Cholangic mucosa with Crohns disease, Cholangic mucosa with ulcerative colitis, Peta1 thymus, Cervix, Cervical adenosquamous carcinoma, Ligament cells, Prostate carcinoma, Bladder carcinoma, Brain oligodendroga; NCI CGAP Sub9 is a substracted cDNA library constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was ligated to an EcoR I adaptor, digested with Not I, and cloned directionally into pT73-Pac vector. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag sequence that is located between the Not I site and the (drr)18 tail. The sequence tags for this library are CTCG, AACG, GGGCC, GGAG, TAGC, TAAGC, ATGG, ACAC, ATCAC. For additional information, contact: Bento Soares, bento-soares@uiowa.edu TAG LIB=UI-H-CO0 TAG TISSUE=Cervical Adenosquamous Carcinoma TAG SBO=CGAG"						
BASE COUNT	194	a	113	c	139	g	221	t
ORIGIN								
Alignment Scores:								
Pred. No.:	3,71e-109	Length:	667					
Score:	1117.00	Matches:	219					
Percent Similarity:	100.00%	Conservative:	0					
Best Local Similarity:	100.00%	Mismatch:	0					
Query Match:	15.20%	Indels:	0					
DB:	12	Gaps:	0					
US-10-020-095-4 (1-1428) x BQ028040 (1-667)								
QY	60	AlGluLeuLeuLeuYsThrAlaSerAsnLeuThrValSerValLeuGluAlaGluGlyVal	79					
Db	665	GCGGAGCTGCTCAAGCAGCATCAAACTCACTGCTCTGCTCGAAGCAGAAAGAGTC	606					
QY	80	PheGluValGlySerPheIysThrLeuThrLeuProSerLeuProLeuAsnSerAlaAsp	99					
Db	605	TTTGAAAAAGGCTCTTTTAAAGCACTTCACTTCCATCCATCAACCTCGAACAGGCGAGT	546					
QY	100	GluIleTyrGluLeuArgValThrGlyValArgThrGlnAspGluIleLeuPheSerAsnSer	119					
Db	545	GAGATTTAAGAGCTACGTGTAAACCGAGCTACCCAGATGAGATTTTATCTCTTAATAGT	466					
QY	120	ThrArgLeuSerPheGluThrLysArgIleSerValPheIleGlnThrAspLysAlaLeu	139					
Db	485	ACCGGCTTATCATTTGAGCCAAAGAAATCTGCTTCATTCAAAACAGCAAGGCTTA	426					
QY	140	TyrLysProLysGluGluValLysPheArgIleValThrLeuPheSerAspPheLysPro	159					
Db	425	TACAAACCAAAACCAAGAGTGAAGTTGGCACTTGTACACTCTTCAGATTTTAAAGCT	366					
QY	160	TyrLysThrSerLeuAsnIleLeuIleLysAspProLysSerAsnLeuIleGlnGlnTrp	179					
Db	365	TACAAACCTCTTTAAACCTTCTCACTTAAGACCCCAATTAATTGATCCAAAGTGG	306					
QY	180	LeuSerGlnGlnSerAspLeuGlyValIleSerLysThrPheGlnLeuSerSerHisPro	199					
Db	305	TTGTCAACAACAAGGATCTTGAAGTCAATTTCCAAACTTTTCAGTATCTCCCATCA	246					
QY	200	IleLeuGlyAspTrpSerIleGlnValGlnValAsnAspGlnThrTyrTyrGlnSerPhe	219					
Db	245	ATACCTTGAGCTGGCTCAATTCAAATTCGAATGATGACGACATATTATTCATATCAATTT	186					

QY 220 GlnValSerGluThrValLeuProLysPheGluValThrLeuGlnThrProLeuThrCys 239  
 Db 185 CAGTTTCAGAAATGATGATTAACCAAAATTGCAATGACTTTCAGACACCATTAATATGT 126  
 QY 240 SerMetAsnSerLysHisLeuAsnGlyThrIleThrAlaLysThrThrGlyLysPro 239  
 Db 125 TCTATGAAATCTTAAGCATTTTAAATGATACCATCAAGCAAGTATACATATGGAGACCA 66  
 QY 260 ValIleGluValAspValThrLeuThrPheLeuProLeuSerPheThrGlyLysValLys 278  
 Db 65 GTGAAAGACAGACGTAAACGCTTACCTTTATCCTTTTGGGGGAAAAA 9  
 RESULT 9  
 AL576908/c 1201 bp mRNA linear EST 01-JUN-2003  
 LOCUS AL576908 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA  
 DEFINITION clone CSOD1081YB18 3-PRIME, mRNA sequence.  
 ACCESSION AL576908  
 VERSION AL576908.2 GI:31315187  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.  
 REFERENCE 1. (bases 1 to 1201)  
 Li, W. B., Gruber, C., Jessee, J. and Polayes, D.  
 Full-length cDNA libraries and normalization  
 Unpublished  
 On Feb 16, 2001 this sequence version replaced gi:12939516.  
 COMMENT Contact: Genoscope  
 Genoscope - Centre National de Sequencage  
 BP 191 91006 Evry cedex - France  
 Email: seqref@genoscope.cns.fr, Web: www.genoscope.cns.fr  
 Library was constructed by Life Technologies, a division of  
 Invitrogen. This sequence belongs to sequence cluster 3578.f for  
 more information about this cluster, see  
 http://www.genoscope.cns.fr/  
 cgi-bin/cluster.cgi?seq=CSOD1081DA09NP1&cluster=3578.f. Contact :  
 Feng Liang Email: fliang@life.techn.com URL: http://fulllength.invitrogen.com/Invitrogen Corporation 1600  
 Faraday Avenue Genoscope sequence ID: CSOD1081DA09NP1.  
 FEATURES  
 source  
 1. location/Qualifiers  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="CSOD1081YB18"  
 /issue\_type="PLACENTA COT 25-NORMALIZED"  
 /clone\_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"  
 /note="1st strand cDNA was primed with a NotI-oligo (dT)  
 primer. Five prime end enriched, double-strand cDNA was  
 digested with Not I and cloned into the Not I and EcoR V  
 sites of the pCMVSPORT 6 vector. Library was normalized."  
 BASE COUNT 344 a 248 c 224 g 321 t 64 others  
 ORIGIN  
 Alignment Scores:  
 Pred. No.: 1.6e-98 Length: 1201  
 Score: 1022.00 Matches: 222  
 Percent Similarity: 88.84% Conservat: 1  
 Best Local Similarity: 88.45% Mismatches: 27  
 Query Match: 13.91% Indels: 5  
 DB: 9 Gaps: 1  
 US-10-020-095-4 (1-1428) x AL576908 (1-1201)  
 QY 1178 PheAlaSerThrGlnAspThrThrValAlaLeuLysAlaLeuSerGluPheAlaLeu 1197  
 Db 1030 TTTCATCTACTCA-GATACACCT--KTGCTTAAGCTCTTCTTTRA-ATTGCAGCCCTA 976  
 QY 1198 MetAsnThrGluValGlyThrAsnIleGlnValThrValThrGlyProSerSerProSerPro 1217  
 Db 1198 MetAsnThrGluValGlyThrAsnIleGlnValThrValThrGlyProSerSerProSerPro 1217

Db 975 ATGAAWACAGAAAG-ACAAATATCCAGATGACCGGCT-AGCTCACCAAGTCTCT 918  
 QY 1218 LeuAlaValAlaGlnProMetAlaValAsnIleSerAlaAsnGlyPheGluPheAlaIle 1237  
 Db 917 CTCGCTGTGGAACAGCCACGCGCATTAATTTCCGCAATGGTTTGGATTGCTATT 858  
 QY 1238 CygGlnLeuAsnValValIyrAsnValLysAlaSerGlySerSerArgArgArgSer 1257  
 Db 857 TGTCACTCAATGTTGTATATATATGAGAGGCTTCGGGCTCTTCAAGACAGCAAGATCT 798  
 QY 1258 IleGlnAsnGlnAlaPheAspLeuAspValAlaValIleGlnLeuLysAspAspLeu 1277  
 Db 797 ATCCAAATCCAAAGAACCTTTGATTGATGATGCTGTAAGAAATAAGATGATCTC 728  
 QY 1278 AsnHleValAspLeuAsnValCysThrSerPheSerGlyProGlyArgSerGlyMetAla 1297  
 Db 737 AATCATGTGATTTGAATGTGTGAACAGCTTTCCGGCCCGGTAGAGTGCATGCT 678  
 QY 1298 LeuMetGluValAsnLeuLeuSerGlyPheMetValProSerGluAlaIleSerLeuSer 1317  
 Db 677 CTTATGGAAGTTAACCTATTAACTGCTTATGCTGCTTCAAGACCAATTTCTTGAGC 618  
 QY 1318 GluThrValLysLysValGluThrAspHisGlyLysLeuAsnLeuThrLeuAspSerVal 1337  
 Db 617 GAGACAGTGAAAGAAAGTGAAATGTGATCATGAAACTCAACCTCATATTGATTCGTA 558  
 QY 1338 AsnGluThrGlnPheCysValAsnIleProAlaValArgAsnPheLysValSerAsnThr 1357  
 Db 557 AATGAACCCAGTTTGTGTATATATNCTGCTGTGAGAACTTTAAGTTCATAATACC 498  
 QY 1358 GlnAspAlaSerValSerIleValAspThrThrGluProArgArgGlnAlaValArgSer 1377  
 Db 497 CAAGAGCTTCAGTGTCCATAGATGATGATTAAGTACAGCAAGAGCGGTGAGAGT 438  
 QY 1378 TyrAsnSerGluValLysLeuSerSerCysAspLeuCysSerAspValGlnGlyCysArg 1397  
 Db 437 AACAACTCTGAATGAACTGCCCCCTGTGACCTTGCAGTATGCCAGVCMSCCGM 378  
 QY 1398 ProCysGluAspGlyValAspGlySerHisHisSerSerValIlePheIlePheCys 1417  
 Db 377 CTTGTGAGAGTATGAGCTTCAVCGMCCATCACTCRCACTGATTTTATTCTGT 318  
 QY 1418 PheLysLeuLeuThrPheMetGluLeuThrPhe 1428  
 Db 317 CTCAGCTTCTGAACCTTATGAACCTTGGCTG 285  
 RESULT 10  
 AK042169 1495 bp mRNA linear HTC 05-DEC-2002  
 LOCUS AK042169  
 DEFINITION Mus musculus 3 days neonate thymus cDNA, RIKEN full-length enriched  
 library, clone:A630065N19 product:weakly similar to  
 alpha-2-macroglobulin (fragments) [limulus polyphemus], full insert  
 sequence.  
 ACCESSION AK042169  
 VERSION AK042169.1 GI:26334978  
 KEYWORDS HTC; CAP trapper.  
 SOURCE Mus musculus (house mouse)  
 ORGANISM Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 REFERENCE 1  
 Carninci, P. and Hayashizaki, Y.  
 High-efficiency full-length cDNA cloning  
 Meth. Enzymol. 303, 19-44 (1999)  
 JOURNAL MEDLINE 99279253  
 PUBMED 10349636  
 REFERENCE 2  
 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,  
 Itoh, M., Kono, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.  
 Normalization and subtraction of cap-trapper-selected cDNAs to  
 prepare full-length cDNA libraries for rapid discovery of new genes  
 JOURNAL Genome Res. 10 (10), 1617-1630 (2000)  
 MEDLINE 20499374





Oy		1267	AapValAlaValLysGluAsnLysAspAspLeuSerHisValAlaAspLeuAsnValCysThr	1286
Db		478	GAGCTCATTTGG---AATTAATGAAGACACACTTAACTCACTCGAATCTGAATGTGTGCACA	534
Oy		1287	SerPheSerGlyProGlyArgSerGlyMetAlaLeuMetGluValAsnLeuLeuSerGly	1306
Db		535	AGTCACCTGGGCTTGAGAAGAGACAGGACGTGGCTTATTGGAAGTGAACCTTCCAGTGGC	594
Oy		1307	PheMetValProSerGluAlaIleSerLeuSerGlyIleThrValLysValGluTyrAsp	1326
Db		595	TTTAGTGTCATCTTCAGATTCATTCCTTGAGTGAACCCTGAAGAAAAGTGAATATGAT	654
Oy		1327	HisGlyLysLeuAsnLeuTyrLeuAspSerValAsnGluIleThrGlnPheCysValAsnIle	1346
Db		655	AACGGAAACCTCAACCTTATTGTAATCTCTGAATGAATGCCAGTTTGCTTAATATT	714
Oy		1347	ProAlaValArgAsnPhelLysValSerAsnThrGlnAspAlaSerValSerIleValAsp	1366
Db		715	CCCCCTGTGAAGAGACTTACAAAGTTTCCAATTTTCAGATGTGTCAGTATCTGTAAATGCAC	774
Oy		1367	TyrtYrGluProArgArgGlnAlaValArgSerTyrAsnSerGluValLysLeuSerSer	1386
Db		775	TACTATGACCAAGAGGACGAGCGAGTGGCAAGCTAACACACCAGGTAAACCTGTCTCCA	834
Oy		1387	CysAspLeuCysSerAspValGlnGlyCysArgProCysGluAspGlyAlaSerGlySer	1406
Db		835	TGTTATCCTCATCTCCAGACACC---AAGTCGAAGTACACACAGGACGAGCCACGAGCTCC	891
Oy		1407	HisHisSerSerServAlIlePheIlePheCysPheLysLeuTyrPheMetGlu	1425
Db		892	CTTGACAGCTTCTTGAGCGCTTCTGCTTTGTGCTCCGCTTGTGTAATTTGTGCAA	948
RESULT 11				
LOCUS	BX386457	978 bp	mRNA	EST 08-MAY-2003
DEFINITION	BX386457 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA			
ACCESSION	clone CS001081YB18 5-PRIME, mRNA sequence.			
VERSION	BX386457			
KEYWORDS	BX386457.1 GI:30457430			
SOURCE	EST.			
ORGANISM	Homo sapiens (human)			
	Homo sapiens			
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
	1 (bases 1 to 978)			
	Ii.W.B., Gruber,C.; Jesse,J. and Polayes,D.			
	Full-length cDNA libraries and normalization			
	Unpublished			
REFERENCE	Contact: Genoscope			
AUTHORS	Genoscope - Centre National de Sequencage			
TITLE	BP 191 91006 EVRY cedex - France			
JOURNAL	Email: segre@genoscope.cns.fr, Web : www.genoscope.cns.fr			
COMMENT	Library was constructed by Life Technologists, a division of			
	Invitrogen. This sequence belongs to sequence cluster 3578.f For			
	more information about this cluster, see			
	http://www.genoscope.cns.fr/			
	cgt-bin/cluster_cgt/seg=CSAI021BI00Plcuster=3578.f. Contact :			
	Feng Liang Email : fliang@life tech.com URL : http://fulllength.invitrogen.com/ Invitrogen Corporation 1600			
	Paradey Avenue Genoscope sequence ID : CSAI021BI00Pl.			

BASE COUNT	374 a	151 c	179 g	220 t	54 others
ORIGIN					
Alignment Scores:					
Pred. No.:	6.09e-97	Length:		978	
Score:	1006.00	Matches:		217	
Percent Similarity:	75.59%	Conservative:		9	
Best Local Similarity:	72.58%	Mismatches:		72	
Query Match:	13.69%	Indels:		2	
DB:	13	Gaps:		0	
US-10-020-095-4 (1-1428) x BXJ6457 (1-978)					
QY	3	GlyProProLeuLeuThrAlaAlaHisLeuLeuCyValCysThrAlaAlaLeuAlaVal	22		
DB	84	GGGCGAACCCTCTCTGAACGAGCAAAATCTCTGGCTGACCGCGCGCTGGCGTG	143		
QY	23	AlaProGlyProArgPheLeuValThrAlaProGlyIleIleArgProGlyGlyAsnVal	42		
DB	144	GCTCCAGGCGCGCTGTTTGCTGTGAGAGCGCAGGAGATATAGAGCAAGAGAAATGTG	203		
QY	43	ThrIleGlyValGluLeuLeuGluHisCysProSerGlnValThrValValaGluLeu	62		
DB	204	ACWATTGGGTGGAGCTCTCTGGAACACTGCMCTTACAGTGAGCTGGAAGGCGGA-SWG	262		
QY	63	LeuIleThrAlaSerAsnLeuThrValSerValLeuGluValaGluGlyValaPheGluLys	82		
DB	263	CTCAAGCAGCAGCAATMAATMAAGTATCAGTAATGAGAAAGAGAGTTTGAATAA	322		
QY	83	GlySerPheIleThrLeuThrLeuProSerLeuProLeuAsnSerAlaAspGluIleTyr	102		
DB	323	AGATCTTTTAAATAAATAAATAATCAAAAAAATAAATAAATAAATAAATAAATAA	382		
QY	103	GluLeuArgValThrGlyArgThrGlnAspGluIleLeuPheSerAsnSerThrArgLeu	122		
DB	383	AACMAAGGTAAAGAGAGTACCCAGGATGARATWTAACTCTAAMAGAACAMGCTA	442		
QY	123	SerPheGluThrIleArgIleSerValPheIleGlnThrAspIleValaLeuTyrLysPro	142		
DB	443	AAATTTGAAACCAAAAAATAAAGAAATMAAAAAAATAAATAAATAAATAAATAA	502		
QY	143	LysGlnGluValaLysPheArgIleValThrLeuPheSerAspPheLysProTyrLysThr	162		
DB	503	AAGCAAGAGTGAAGTTTGGCTTGTACCTCTTAATAATTAAGCATTAACAAAAA	562		
QY	163	SerLeuAsnIleLeuIleLysAspProLysSerAsnLeuIleGlnIleProLysSerGln	182		
DB	563	ACTTAAACATTATCATTAAGAAACCAATTAATTTGATCCAAACAGTGGTGTCAAA	622		
QY	183	GlnSerAspLeuGlyValIleSerIleThrPheGlnLeuSerSerHisProIleLeuGly	202		
DB	623	CAAGAGATCTGGAGTCAATTCCAAACTTTGACGATCTTCCATCCAAATACCTGGT	682		
QY	203	AspThrSerIleGlnValaGlnValaAsnAspGlnThrTyrGlnIleSerPheGlnValaSer	222		
DB	683	GACGGCTATTAAGTCAAGTGAATGACGACATATATATCAATCATTTCCAGTTTCA	742		
QY	223	GluTyrValLeuProLysPheGluValaThrLeuGlnThrProLeuTyrCysSerMetAsn	242		
DB	743	GAATATGATTAACAAATTAATTAATGATCTTTGCGAGACACATTAATTTGTTATGAAT	802		
QY	243	SerLysHisLeuAsnGlyThrIleThrAlaLysTyrThrTyrGlyLysProValLysGly	262		
DB	803	TCTAACCATTTAAATGATCAATTCATTCCAAAASATATAGATTSTAAGTAAAGYAAATCA	862		
QY	263	AspValThrLeuThrPheLeuProLeuSerPheTyrPglLys-LysLysAsnIleThrLys	282		
DB	863	GACTTACCTTACATTTTACCTTTTACCTTTCCKTGTGGGSAAKAKAKAAAAATATTACAA	922		
QY	282	gThrPheLysIleAsnGlySerAlaAsnPheSerPheAsnAspGluGluMetLys	300		
DB	923	AACATTTAAGATTAATGATCTGCAAAATTTCTCTTTTATATGATGAAGATGAAA	977		

RESULT 12  
Bg173799 836 bp mRNA linear EST 06-FEB-2001  
LOCUS 60233853F1 NCI\_CGAP\_Mam1 Mus musculus cDNA clone IMAGE:4457172 5',  
DEFINITION mRNA sequence.  
ACCESSION Bg173799.1 GI:12680502  
VERSION Bg173799.1  
KEYWORDS EST.  
SOURCE Mus musculus (house mouse)  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
REFERENCE 1 (bases 1 to 836)  
AUTHORS NIH-MGC <http://mgi.nci.nih.gov/>.  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: [cgapbs-remail.nih.gov](mailto:cgapbs-remail.nih.gov)  
Tissue Procurement: Gilbert Smith, Ph.D.  
cDNA Library Preparation: Life Technologies, Inc.  
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LNL at:  
<http://image.llnl.gov>  
Plate: LLM10253 row: 1 column: 13  
High quality sequence stop: 694.  
Location/Qualifiers

FEATURES  
source  
1..836  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="FVB/N"  
/db\_xref="taxon:10090"  
/clone="IMAGE:4457172"  
/tissue\_type="tumor, biopsy sample"  
/dev\_stage="10 months, virgin"  
/lab\_host="DH10B"  
/note="Organ: mammary; Vector: pCMV-SPORT6; Site: 1; Salt; Site: 2; NotI; Cloned unidirectionally. Primer: Oligo dT. Library constructed by Life Technologies. Investigator providing samples: Gilbert Smith, NIH"

BASE COUNT 228 a 179 c 202 g 227 t  
ORIGIN

Alignment Scores:  
Pred. No.: 3.37e-94 Length: 836  
Score: 979.50 Matches: 200  
Percent Similarity: 87.55% Conservative: 25  
Best Local Similarity: 77.82% Mismatches: 30  
Query Match: 13.33% Indels: 4  
DB: 10 Gaps: 1

US-10-020-095-4 (1-1428) x Bg173799 (1-836)

QY 942 LysLysGlnLeuThrAspAsnLeuYsgLysAlaLeuSerPheMetArgInglYtyr 961  
Db 1 CAGAAACAGCTGACGTTAATTAAAGAAAGCCCTTCAATATATAGGCAAGCTTAC 60  
QY 962 GlnArgGluLeuLeuTyrGlnArgGluAspGlySerPheSerAlaPheGlyAsnTyrAsp 981  
Db 61 CAAAGGAGGCTTCTATACAGAGGAGATGCTCTTCAAGCTTTGGGAGCAATTGAC 120  
QY 982 ProSerGlySerThrTyrLeuSerAlaPheValLeuArgCysPheLeuGluAlaAspPro 1001  
Db 121 TCTTTCGAGAGACACTGCTGTACACATTTGTTTAAATGCTTTCTCGAAGCTGATTAC 180  
QY 1002 TyrIleAspIleAspGlnAsnValLeuHisArgThrTyrThrTyrLeuLysGlyHisGln 1021  
Db 181 TATATAGATATTGATCAGATGTGTATACCAAAACATTAACCTTGCTTAATGCAATAG 240  
QY 1022 LysSerAsnGlyGluPheTyrAspProGlyArgValIleHisSerGluLeuGlnGlyGly 1041

Db 241 AATTCATAGTGAATTTGGAGCCAGAAAGACTGATTCACTGATCACTTCAAGTGGC 300  
QY 1042 AsnLysSerProValThrLeuThrAlaTyrIleValThrSerLeuLeuGlyTyrArgLys 1061  
Db 301 ACCAAAGCCAGTAACCTTACGCTTATATGTGACTTGTGCTGGGACACAAAAG 360  
QY 1062 TyrGlnProAsnIleAspValGlnGluSerIleHisPheLeuGluSerGluPheSerArg 1081  
Db 361 TATAGCCTATATCGATGATCAAGACATCAAGTTCATCAAGTTCATCAAGTTCATCAAG 419  
QY 1082 GlyIleSerAspAsnTyrThrLeuAlaLeuIleThrTyrAlaLeuSerValGlySer 1101  
Db 420 GGAATTCGAGCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 479  
QY 1102 ProLysAlaLysGluAlaLeuAsnMetLeuThrTyrArgIleGluGlnGlyGlyMet 1121  
Db 480 CTTAAAGAGAGAGGCTTGAACCTCTGATGCAAGCATGAAAGAGAGAGACACA 539  
QY 1122 GlnPheTyrValSerSerGluSerLysLeuSerAspSerTyrGlnProArgSerLeuAsp 1141  
Db 540 CAGTTCTGTGTATACAGACCTGCGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 599  
QY 1142 IleGluValAlaAlaTyrAlaLeuLeuSerHisPheLeuGlnPheGlnThrSerGlyGly 1161  
Db 600 ATCGAAATCGCTGTTATGCACTGTGTGCGCACCTGTGAC--CATGTCTTGAAGA 656  
QY 1162 IleProIleMetArgTyrPheLeuSerArgGlnArgAsnSerLeuGlyGlyPheAlaSerTh 1181  
Db 657 ATCCGGTTATGAAGTGGCTTCATCAGCAAGAAAGAGCTGGAGG-TTTGTAATCCAC 715  
QY 1181 rGlnAspThrThrValAlaLeuLysAlaLeuSerGluPheAlaAlaLeu 1197  
Db 716 TCAGATATTGTTGTGGGCTTGAAGCAATTTTGAATTTTCCGCTTG 764

RESULT 13  
B0624356/c 638 bp mRNA linear EST 23-SEP-2002  
LOCUS UI-H-FG1-bgj-1-22-0-UI-s1 NCI\_CGAP\_FG1 Homo sapiens cDNA clone  
DEFINITION UI-H-FG1-bgj-1-22-0-UI 3', mRNA sequence.  
ACCESSION B0624356  
VERSION B0624356.1 GI:23290571  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 638)  
AUTHORS NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.  
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
Tumor Gene Index  
JOURNAL Unpublished  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: [cgapbs-remail.nih.gov](mailto:cgapbs-remail.nih.gov)  
Tissue Procurement: James Martin  
cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa  
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa  
Clone Distribution: Clone distribution information can be obtained  
from Dr. M. Bento Soares, [bento-soares@uiowa.edu](mailto:bento-soares@uiowa.edu)  
Seq primer: M13 FORWARD  
POLY-A=Yes.  
Location/Qualifiers

FEATURES  
source  
1..638  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="UI-H-FG1-bgj-1-22-0-UI"  
/tissue\_type="Cell lines"  
/dev\_stage="Adult"  
/lab\_host="DH10B (Life Technologies)"  
/clone\_lib="NCI\_CGAP FG1"  
/note="Organ: Enchondroma; Vector: pT73-Pac (Pharmacia)  
with a modified polylinker; Site\_1: Bcor I; Site\_2: Not I;



prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. cDNA went through one round of normalization to Rot = 10.0 and subtraction to Rot = 100.0. Second strand cDNA was prepared with the primer adapter of sequence [5' GAGAGAGAGATTCGATTATTAATATCCCCCCCCC 3']. cDNA was cloned into the XhoI and BamHI sites. Vector: a modified phagescript KS(+) after bulk excision from lambda FLC I."

BASE COUNT 195 a 118 c 140 g 194 t 1 others  
ORIGIN

## Alignment Scores:

Pred. No.: 3,12e-90 Length: 648  
Score: 941.00 Matches: 178  
Percent Similarity: 92.09% Conservative: 20  
Best Local Similarity: 82.79% Mismatches: 17  
Query Match: 12.81% Indels: 0  
DB: 10 Gaps: 0

US-10-020-095-4 (1-1428) x BB613975 (1-648)

QY 862 GlnSerIleuLeuAspLeuThrAspAsnArgLeuGlnSerThrLeuYsThrLeuSer 881  
DB 2 AAATCTGTTATGGATCTGACAGATACCAAGTAAGACCAATCTATGAGA 61  
QY 882 PheSerPheProProAsnThrValThrGlySerGluArgValGlnIleThrAlaIleGly 901  
DB 62 TTTCTCTTCTCTCGATGAGGATCGATGCGAGTAAGTAAGTACAGATCAAGCAATGGA 121  
QY 902 AspValIleuGlyProSerIleuAsnGlyLeuAlaSerIleuIleArgMetProTyrGlyCys 921  
DB 122 GATATCCTTGTTCCCTCCATCAATGCTTATCTTCACTGATCGAGATGCTTATGAGAT 181  
QY 922 GlyIleuGlnAsnMetIleAsnPheAlaProAsnIleTyrIleuAspTyrIleuThrIys 941  
DB 182 GGTGAACAGAACATGATATATTTGCTCCAAATATTACATTTGGATTATCTGACTAAA 241  
QY 942 LysIleuGlnLeuThrAspAsnLeuGlyValAlaLeuSerPheMetArgGlnGlyTyr 961  
DB 242 CAGAAACAGCTGACGATTATTAAGAAAGCCCTTCAATATGAGGCAAGGTTAC 301  
QY 962 GlnArgGluLeuLeuTyrGlnArgIleuAspGlySerPheSerAlaPheGlyAsnTyrAsp 981  
DB 302 CAAAGGAGCTTCTCATCAGAGGAGAGAGTGGCTCTTCAAGCTTTGGGACATGAC 361  
QY 982 ProSerGlySerThrTyrLeuSerAlaPheValIleuArgCysPheLeuGlnAlaAspPro 1001  
DB 362 TCTCTCGGAGACCTTGCTGACGATTGTTTAAAGATGCTTCTGGAAGCTGATTAC 421  
QY 1002 TyrIleAspIleAspGlnAsnValIleuHisArgThrTyrThrTyrLeuIysGlnIle 1021  
DB 422 TATATAGATATGATGAGATGTTTACACAGAACTATATCTGGCTTATATGACATGAG 481  
QY 1022 LysSerAsnGlyIleuPheTyrAspProGlyArgValIleHisSerGluLeuGlnGlyIle 1041  
DB 482 AAATTCATAGTGAATTTGGAGCCAGAGACGATTCACAGTCACTTAAGGTGC 541  
QY 1042 AsnLysSerProValThrLeuThrAlaTyrIleValThrSerLeuLeuGlyTyrArgIys 1061  
DB 542 ACCAAAGCCAGTAACTTACGGGCTATATTTGATCTTGTCTGGGATANCAAAAG 601  
QY 1062 TyrGlnProAsnIleAspValGlnGlySerIleHisPheLeuGln 1076  
DB 602 TATCAGCCTAATATCATGATGATACAGACTCAATCAAGTTTGGAA 646

## RESULT 15

LOCUS AA418644 634 bp mRNA linear EST 16-OCT-1997  
DEFINITION ZV93F07.r1 Soares NHMPu S1 Homo sapiens cDNA clone IMAGE:767365 5' similar to TR:G534873 G534873 ALPHA-2-MACROGLOBULIN, ;, mRNA  
ACCESSION AA418644

VERSION AA418644.1 GI:2080463  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo. 1 (bases 1 to 634)

AUTHORS Hillier, L., Allen, M., Bowles, L., Dubuque, T., Geisel, G., Jost, S., Kuuba, T., Lacy, M., Le, N., Lennon, G., Marra, M., Martin, J., Moore, B., Schellenberg, K., Stepien, M., Tan, F., Theising, B., White, Y., Wyllie, T., Waterston, R., and Wilson, R.

TITLE WashU-Merck EST Project 1997  
JOURNAL Unpublished  
COMMENT Contact: Wilson RK  
Washington University School of Medicine  
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
Tel: 314 286 1800  
Fax: 314 286 1810  
Email: est@watson.wustl.edu

This clone is available royalty-free through LNL; contact the IMAGE Consortium (info@image.llnl.gov) for further information.  
Seq primer: -28m13 rev2 ET from Amersham  
High quality sequence stop: 494.  
Location/Qualifiers  
1..634  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:767365"  
/tissue\_type="Pooled human melanocyte, fetal heart, and pregnant uterus"  
/lab\_host="DH10B"  
/clone\_lib="Soares NHMPu S1"  
/note="Organ: mixed (see below); Vector: pTTT3D-pac (pharmacia) with a modified polylinker; Site 1: Not 1; Site 2: Eco RI; Equal amounts of plasmid DNA from three normalized libraries (melanocyte 2NBH, pregnant uterus NBHPU, and fetal heart NBH19W) were mixed, and ss circles were made in vitro. Following HAP purification, this DNA was used as tracer in a subtractive hybridization reaction. The driver was PCR-amplified cDNAs from pools of 5,000 clones made from the same 3 libraries. The pools consisted of I.M.A.G.E. clones 260232-265223, 340488-345479, and 484488-489479."

## FEATURES

## source

BASE COUNT 174 a 109 c 139 g 212 t  
ORIGIN  
Alignment Scores:  
Pred. No.: 3,99e-87 Length: 634  
Score: 912.00 Matches: 174  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 12.41% Indels: 0  
DB: 9 Gaps: 0

US-10-020-095-4 (1-1428) x AA418644 (1-634)

QY 1255 ArgArgSerIleGlnAsnGlnAlaPheAspLeuAspValAlaValIleGlnAsnIys 1274  
DB 1 CGAAGATCTATCCAAATCAAGAACCTTTGATTAGATGTTGCTGTAAGAAATATAA 60  
QY 1275 AspAspLeuAsnHisValAspLeuAsnValCysThrSerPheSerGlyProGlyArgSer 1294  
DB 61 GATGATCTCAATCATGAGGATTTGAATGTGTACCAAGCTTTTGGGCGCGGTAGAGGT 120  
QY 1295 GlyMetAlaLeuMetGluValAsnLeuLeuSerGlyPheMetValProSerGluAlaIle 1314  
DB 121 GGCAATGCTCTTATGGAAGTTAACTTAAGTGGCTTATAGTGCGCTTCAAGAACATT 180  
QY 1315 SerLeuSerGluThrValIleValIleValGluTyrAspHisGlyIleLeuAsnLeuTyrIleu 1334  
DB 181 TCTCTGACGACAGACGATGAAAGAGTGAATATGATCATGAGAAATCAACCTCTATTTA 240

```
Qy 1335 AspSerValAsnGluThrGlnPheCysValAsnIleProAlaValArgAsnPhelysVal 1354
    |||||
Db 241 GATTCTGTAATGAAGAACCCAGTTTGTGTATATCTCTGCTGAGAGAACTTTAAAGTT 300
    |||||
Qy 1355 SerAsnThrGlnAspAlaSerValSerIleValAspTyrTyrGluProArgArgGlnAla 1374
    |||||
Db 301 TCAATATCCCAAGATGCTTCCAGTGTCTCAATGATTTACTATGAGCCAGAGAGACAGGCG 360
    |||||
Qy 1375 ValArgSerTyrAsnSerGluValIlyleuSerSerCysAspLeuCysSerAspValGln 1394
    |||||
Db 361 GTGAGAGATTACACTGTAAGTGAAGCTGTCTCTGTGACCTTTGAGATGATGTCAG 420
    |||||
Qy 1395 GlyCysArgProCysGluAspGlyAlaSerGlySerHisHisSerSerValIlePhe 1414
    |||||
Db 421 GGCTGCCGCTCTGTGAGGATGAGAGCTTCCATCATCACTCTTCAGTCATTTT 480
    |||||
Qy 1415 IlePheCysPheLysLeuLeuTyrPheMetGluLeuTyrPhe 1428
    |||||
Db 481 ATTTCGTTCAGAGCTTGTGACTTATGSACTTGGCTG 522
    |||||
```

Search completed: January 17, 2004, 02:52:00  
Job time : 5723 secs





PA (UYMC-) UNIV MCGILL.  
 XX  
 PI Philip A, Tam B;  
 XX  
 DR WPI, 2003-093100/08.  
 DR N-PSDB; AAD49434.  
 XX  
 PT Novel transforming growth factor (TGF)-beta 1 binding reagent which  
 PT comprises r150 protein which acts as accessory receptor of TGF-beta,  
 PT useful for negatively modulating TGF-beta activity, and thus for  
 PT treating cancer -  
 XX  
 PS Claim 1; Page 94-99; 127pp; English.  
 XX  
 CC The invention relates to novel transforming growth factor (TGF)-beta1  
 CC binding reagent which comprises a glycosylphosphatidylinositol (GPI)-  
 CC anchored TGF-beta1 binding protein referred to as r150 which acts as  
 CC an accessory receptor of TGF-beta. The invention is used for negatively  
 CC modulating TGF-beta activity, and thus for treating conditions  
 CC characterised by overproduction of TGF-beta, such as cancer. Antisense  
 CC molecules of the invention are used for increasing TGF-beta availability  
 CC and increase graft success. The present sequence is human r150 protein.  
 XX  
 SQ Sequence 1428 AA;  
 Query Match 99.8%; Score 7331; DB 24; Length 1428;  
 Best Local Similarity 99.8%; Pred. No. 0;  
 Matches 1425; Conservative 1; Mismatches 2; Indels 0; Gaps 0;  
 QY 1 MCGPELLTAHLCTAALAVADGPPPLVTAAGIIRPGNVITGVELLHCPQVTVKA 60  
 DB 1 MCGPELLTAHLCTAALAVADGPPPLVTAAGIIRPGNVITGVELLHCPQVTVKA 60  
 QY 61 ELKTAANLVSYLEAGVPEKSGFKTLTLPLNADETIELKRVNGRQDILLFNSNT 120  
 DB 61 ELKTAANLVSYLEAGVPEKSGFKTLTLPLNADETIELKRVNGRQDILLFNSNT 120  
 QY 61 ELKTAANLVSYLEAGVPEKSGFKTLTLPLNADETIELKRVNGRQDILLFNSNT 120  
 DB 61 ELKTAANLVSYLEAGVPEKSGFKTLTLPLNADETIELKRVNGRQDILLFNSNT 120  
 QY 121 RLSEETKRIISVFIOTDALKYKQEVKFRIVTLFSDPKPKYKTSNIIILKPKSNLIQOML 180  
 DB 121 RLSEETKRIISVFIOTDALKYKQEVKFRIVTLFSDPKPKYKTSNIIILKPKSNLIQOML 180  
 QY 121 RLSEETKRIISVFIOTDALKYKQEVKFRIVTLFSDPKPKYKTSNIIILKPKSNLIQOML 180  
 DB 121 RLSEETKRIISVFIOTDALKYKQEVKFRIVTLFSDPKPKYKTSNIIILKPKSNLIQOML 180  
 QY 181 SQGSDLVISKTFLQSSHPILGMSIQVQNDQTYGQFVSEVYLKPFVTLQTPLYCS 240  
 DB 181 SQGSDLVISKTFLQSSHPILGMSIQVQNDQTYGQFVSEVYLKPFVTLQTPLYCS 240  
 QY 181 SQGSDLVISKTFLQSSHPILGMSIQVQNDQTYGQFVSEVYLKPFVTLQTPLYCS 240  
 DB 181 SQGSDLVISKTFLQSSHPILGMSIQVQNDQTYGQFVSEVYLKPFVTLQTPLYCS 240  
 QY 241 MNSGHLNGITTAKTYYGKPVGDTLTLPLSPFWGKKKNTTKTKINGSANFSNDEEMK 300  
 DB 241 MNSGHLNGITTAKTYYGKPVGDTLTLPLSPFWGKKKNTTKTKINGSANFSNDEEMK 300  
 QY 241 MNSGHLNGITTAKTYYGKPVGDTLTLPLSPFWGKKKNTTKTKINGSANFSNDEEMK 300  
 DB 241 MNSGHLNGITTAKTYYGKPVGDTLTLPLSPFWGKKKNTTKTKINGSANFSNDEEMK 300  
 QY 301 NVMDSSNGLSEYLLDSSPGVEILITVYESVTGISRNVSTVFPKQHDYIIEFDYTTVL 360  
 DB 301 NVMDSSNGLSEYLLDSSPGVEILITVYESVTGISRNVSTVFPKQHDYIIEFDYTTVL 360  
 QY 301 NVMDSSNGLSEYLLDSSPGVEILITVYESVTGISRNVSTVFPKQHDYIIEFDYTTVL 360  
 DB 301 NVMDSSNGLSEYLLDSSPGVEILITVYESVTGISRNVSTVFPKQHDYIIEFDYTTVL 360  
 QY 361 KPSLNFATVAVKTRADGNQLTLLEERRNNVITVTRQRYTEWSSGNSGNQMEAVQKINY 420  
 DB 361 KPSLNFATVAVKTRADGNQLTLLEERRNNVITVTRQRYTEWSSGNSGNQMEAVQKINY 420  
 QY 361 KPSLNFATVAVKTRADGNQLTLLEERRNNVITVTRQRYTEWSSGNSGNQMEAVQKINY 420  
 DB 361 KPSLNFATVAVKTRADGNQLTLLEERRNNVITVTRQRYTEWSSGNSGNQMEAVQKINY 420  
 QY 421 TVPQSGTFKIEFFPLEBSSSELQKAYFLGSKSMAVHSLFSPSKTYIQLKTRDENIKVG 480  
 DB 421 TVPQSGTFKIEFFPLEBSSSELQKAYFLGSKSMAVHSLFSPSKTYIQLKTRDENIKVG 480  
 QY 421 TVPQSGTFKIEFFPLEBSSSELQKAYFLGSKSMAVHSLFSPSKTYIQLKTRDENIKVG 480  
 DB 421 TVPQSGTFKIEFFPLEBSSSELQKAYFLGSKSMAVHSLFSPSKTYIQLKTRDENIKVG 480  
 QY 481 SPFELVYSGNKRLELSVMVVRGQIVAVGKONSTWMSLTPENSWTPKACIVYIYIEDDG 540  
 DB 481 SPFELVYSGNKRLELSVMVVRGQIVAVGKONSTWMSLTPENSWTPKACIVYIYIEDDG 540  
 QY 481 SPFELVYSGNKRLELSVMVVRGQIVAVGKONSTWMSLTPENSWTPKACIVYIYIEDDG 540  
 DB 481 SPFELVYSGNKRLELSVMVVRGQIVAVGKONSTWMSLTPENSWTPKACIVYIYIEDDG 540  
 QY 541 EIIISDVLPKIPQVLFKQKIKLYMSKVAEPESEKSLRISVTPQPSIYIGIYAVDKSVMIMN 600  
 DB 541 EIIISDVLPKIPQVLFKQKIKLYMSKVAEPESEKSLRISVTPQPSIYIGIYAVDKSVMIMN 600  
 QY 541 EIIISDVLPKIPQVLFKQKIKLYMSKVAEPESEKSLRISVTPQPSIYIGIYAVDKSVMIMN 600  
 DB 541 EIIISDVLPKIPQVLFKQKIKLYMSKVAEPESEKSLRISVTPQPSIYIGIYAVDKSVMIMN 600  
 QY 601 ASNDITMENYVHELELNTGYLLGMFMNSFAVFOECGLMVLTDANLTKDYIDGYDNAEY 660  
 DB 601 ASNDITMENYVHELELNTGYLLGMFMNSFAVFOECGLMVLTDANLTKDYIDGYDNAEY 660  
 QY 601 ASNDITMENYVHELELNTGYLLGMFMNSFAVFOECGLMVLTDANLTKDYIDGYDNAEY 660  
 DB 601 ASNDITMENYVHELELNTGYLLGMFMNSFAVFOECGLMVLTDANLTKDYIDGYDNAEY 660

QY 661 AERFMEBNEGHIVDHDPSLGSSPHVRKHPETWIMLDITNMGYRIQEEFVTVPSDITSW 720  
 DB 661 AERFMEBNEGHIVDHDPSLGSSPHVRKHPETWIMLDITNMGYRIQEEFVTVPSDITSW 720  
 QY 721 VATEGVISEDLGLTTPPELOAFQPFIFLNLPSVIRGEFALEITFNILKATEV 780  
 DB 721 VATEGVISEDLGLTTPPELOAFQPFIFLNLPSVIRGEFALEITFNILKATEV 780  
 QY 781 KVIIEKSDKFDIMTSESEINATGHQOTLLVPSBDGATVPPIPTHGEIPITVTLSP 840  
 DB 781 KVIIEKSDKFDIMTSESEINATGHQOTLLVPSBDGATVPPIPTHGEIPITVTLSP 840  
 QY 841 ASDAVTOMILVKAEGIEKSYQSILLDTNRLQSTLKTLSFSPPNVTGSSRVOITAI 900  
 DB 841 ASDAVTOMILVKAEGIEKSYQSILLDTNRLQSTLKTLSFSPPNVTGSSRVOITAI 900  
 QY 901 GDTLGSINGLASLIMPFGCGQNMNINPAPNIYIIDLYTKKQOLDNLKELSPKROG 960  
 DB 901 GDTLGSINGLASLIMPFGCGQNMNINPAPNIYIIDLYTKKQOLDNLKELSPKROG 960  
 QY 961 YORELLYOREDGSFSAFGNYPDPSGTWLSAFVLRCELPADPYIDIDQNVLHRTYWLK 1020  
 DB 961 YORELLYOREDGSFSAFGNYPDPSGTWLSAFVLRCELPADPYIDIDQNVLHRTYWLK 1020  
 QY 1021 QKNGEFMDPGRVIHSELQGNKSPVTLTAYIVTSLIGRYKQPNIDVOESIHFLESEFS 1080  
 DB 1021 QKNGEFMDPGRVIHSELQGNKSPVTLTAYIVTSLIGRYKQPNIDVOESIHFLESEFS 1080  
 QY 1081 RGISDNYTTLALITYALSSVGSPPAKKALNMLTTRABQEGQMOWWSESLSDSQPRSL 1140  
 DB 1081 RGISDNYTTLALITYALSSVGSPPAKKALNMLTTRABQEGQMOWWSESLSDSQPRSL 1140  
 QY 1141 DIEVAAYALLSHFLQOTSSEGIPIKMWLSRORSLSGFASTODTPTAALKALSEFALMNT 1200  
 DB 1141 DIEVAAYALLSHFLQOTSSEGIPIKMWLSRORSLSGFASTODTPTAALKALSEFALMNT 1200  
 QY 1201 ERTNIQVTVTPSSPSPPLAVVQMAVNIISANGFPAICQNLVVYVYKASSSRRRSION 1260  
 DB 1201 ERTNIQVTVTPSSPSPPLAVVQMAVNIISANGFPAICQNLVVYVYKASSSRRRSION 1260  
 QY 1261 QEAFDIDVAVKENKDLNHYDLNVCISFGSGRSGMALMEVNLISGFMVPSSEALISSEY 1320  
 DB 1261 QEAFDIDVAVKENKDLNHYDLNVCISFGSGRSGMALMEVNLISGFMVPSSEALISSEY 1320  
 QY 1321 KKYEDHGKLNLDLSVNETQFCVNI PAVRNPVSTODASVSIIVYIEPRROAVRSYNS 1380  
 DB 1321 KKYEDHGKLNLDLSVNETQFCVNI PAVRNPVSTODASVSIIVYIEPRROAVRSYNS 1380  
 QY 1381 EYKLSSCDLCSVDQGRPCEDGASGSHHSSVIFICFKLLYEMELML 1428  
 DB 1381 EYKLSSCDLCSVDQGRPCEDGASGSHHSSVIFICFKLLYEMELML 1428

RESULT 2  
 ABB82165  
 ID ABB82165 standard; Protein; 1445 AA.  
 XX  
 AC ABB82165;  
 XX  
 DT 23-DEC-2002 (first entry)  
 XX  
 XX Human CD109 K1 protein.  
 XX  
 DE Human CD109 K1 protein.  
 XX  
 KW CD109; alpha2 macroglobulin; thioester; cerebroprotective; cardiant;  
 KW immunosuppressive; haemostatic; anticoagulant; thrombolytic; human;  
 KW cardiovascular; vasotropic; gene therapy; CD109 K1.  
 OS Homo sapiens.  
 XX  
 XX  
 XX W0200270696-A2.  
 XX  
 XX  
 PD 12-SEP-2002.  
 XX

PF 07-MAR-2002; 2002WO-CA00292.  
 XX 07-MAR-2001; 2001US-273814P.  
 XX (SCHU/) SCHUH A.  
 PA (SUTH/) SUTHERLAND R D.  
 XX Schuh A, Sutherland RD;  
 PI WPI; 2002-713450/77.  
 DR N-PSDB; ABQ79964.  
 XX  
 PT New CD109 nucleic acids and polypeptides, useful in gene therapy,  
 PT particularly for treating strokes, myocardial infarctions, thrombosis,  
 PT thrombocytopenia, autoimmune diseases, or organ or bone marrow  
 PT transplantation -  
 XX  
 PS Claim 21; Fig 3a; 156pp; English.  
 XX  
 CC The invention relates to isolated nucleic acid molecules encoding CD109  
 CC polypeptides. These nucleic acid molecules include the human CDNA  
 CC sequences comprising CD109 K1, CD109 K1-H7, CD109 K15 or their variants.  
 CC CD109 is a novel member of the alpha2 macroglobulin (alpha2M/C3, C4, C5  
 CC family of thioester-containing proteins. The CD109 polypeptides can be  
 CC expressed by standard recombinant methodology. The CD109 nucleic acid,  
 CC CD109 polypeptide or its mimetic, or the anti-CD109 antibody is useful  
 CC for treating or detecting a disease or disorder, e.g. conditions  
 CC associated with endothelial activation, platelet activation, activation  
 CC of the coagulation or fibrinolytic systems, activation of T lymphocytes  
 CC and of the complement system, quantitative or qualitative abnormalities  
 CC of platelet function, increased or impaired platelet aggregation and  
 CC activation, increased or impaired activation of the coagulation and/or  
 CC fibrinolytic systems, or impaired or increased immune activation. These  
 CC are also useful for treating cardiovascular disorders, stroke, myocardial  
 CC infection, thrombosis, embolism, peripheral vascular disease,  
 CC thrombocytopenia, thrombocythemia, autoimmune diseases, organ  
 CC transplantation, or bone marrow transplantation. The CD109 nucleic acids  
 CC or their antisense nucleotide sequence are useful in gene therapy for  
 CC treating these conditions. The present sequence represents the human  
 CC CD109 K1 sequence.  
 CC  
 SQ Sequence 1445 AA;  
 Query Match 99.7%; Score 7326.5; DB 23; Length 1445;  
 Best Local Similarity 98.8%; Pred. No. 0;  
 Matches 1427; Conservative 1; Mismatches 0; Indels 17; Gaps 1;

DB 361 KPSLNFATVKTADGNQLTEERRNNVITVTGRNTEYWSGNSGNQKMEAVOKINY 420  
 QY 421 TVPQSGTFKIEPPIEDSSELQKAYFLGSKSSMAVHSLFKSPSKTYIQLKTRDENIKVG 480  
 DB 421 TVPQSGTFKIEPPIEDSSELQKAYFLGSKSSMAVHSLFKSPSKTYIQLKTRDENIKVG 480  
 QY 481 SPFELVSGNKKLKELSYNNVSRGQLVAVGKNSYTFSLTPENSWTPKACVIVYIYEDDG 540  
 DB 481 SPFELVSGNKKLKELSYNNVSRGQLVAVGKNSYTFSLTPENSWTPKACVIVYIYEDDG 540  
 QY 541 EIIIDVLKIPVQVLRKNIKILYMSKYKAPSPSKVSLRISYTOPDSIVGIVANDKSVNLNN 600  
 DB 541 EIIIDVLKIPVQVLRKNIKILYMSKYKAPSPSKVSLRISYTOPDSIVGIVANDKSVNLNN 600  
 QY 541 EIIIDVLKIPVQVLRKNIKILYMSKYKAPSPSKVSLRISYTOPDSIVGIVANDKSVNLNN 600  
 DB 601 ASNDITMENVHELLEYNTGYLLGEMNNSFAVFOCGMLVLDANLTKOYIGGVYDNAEY 660  
 QY 601 ASNDITMENVHELLEYNTGYLLGEMNNSFAVFOCGMLVLDANLTKOYIGGVYDNAEY 660  
 DB 601 ASNDITMENVHELLEYNTGYLLGEMNNSFAVFOCGMLVLDANLTKOYIGGVYDNAEY 660  
 QY 661 AERFMEENEGHIVDIDHDFSLGSSPHVRKHPETWTWLDTNNGYRIYQEEVTVPSITSM 720  
 DB 661 AERFMEENEGHIVDIDHDFSLGSSPHVRKHPETWTWLDTNNGYRIYQEEVTVPSITSM 720  
 QY 721 VATGVVISDGLGLTTTPVELQAPQPFIFLNLPSYIRGSEFALETIYNYLKDAREV 780  
 DB 721 VATGVVISDGLGLTTTPVELQAPQPFIFLNLPSYIRGSEFALETIYNYLKDAREV 780  
 QY 781 KVIIEKSKPDIIMTSSSEINATGHQOTLLVSEBDATVLPFRPHLGEIPTVYALAPT 840  
 DB 781 KVIIEKSKPDIIMTSSSEINATGHQOTLLVSEBDATVLPFRPHLGEIPTVYALAPT 840  
 QY 841 ASDAVTOMILVABEIGIEKYSQSILDLTDNRLQSTLKTLSFPFNTVTSERVOITAI 900  
 DB 841 ASDAVTOMILVABEIGIEKYSQSILDLTDNRLQSTLKTLSFPFNTVTSERVOITAI 900  
 QY 901 GVLTPSINGLASLIRMPYCGEQQMNPAPYIYLDYTKKKQITDKLKEKALSPMROG 960  
 DB 901 GVLTPSINGLASLIRMPYCGEQQMNPAPYIYLDYTKKKQITDKLKEKALSPMROG 960  
 QY 961 YORELLYOREDSFAPGNYPDSGSTWLSAFLRCEFLADPVIDIDQVLAHTYTWLKGH 1020  
 DB 961 YORELLYOREDSFAPGNYPDSGSTWLSAFLRCEFLADPVIDIDQVLAHTYTWLKGH 1020  
 QY 1021 QKSNGEFMDPGRAVHISELQGGNKSPTVLTAYIVTSLGYRKYQPINIDVQESIHFLSEFS 1080  
 DB 1021 QKSNGEFMDPGRAVHISELQGGNKSPTVLTAYIVTSLGYRKYQPINIDVQESIHFLSEFS 1080  
 QY 1081 RGISDNYTLALITVALSSVSGSPKAKEALNMLTWRAOEGGQGFVWSSEKLSDSWQPSL 1140  
 DB 1081 RGISDNYTLALITVALSSVSGSPKAKEALNMLTWRAOEGGQGFVWSSEKLSDSWQPSL 1140  
 QY 1141 DIEVAAYALLSHFLOPOTSEGIPIRWLRSQNSLIGSFASVODTVALKALSEPALNMT 1200  
 DB 1141 DIEVAAYALLSHFLOPOTSEGIPIRWLRSQNSLIGSFASVODTVALKALSEPALNMT 1200  
 QY 1201 ERTNIQVTVGSPSSP-----LAVQPMANVINSANGFGAICOLNV 1243  
 DB 1201 ERTNIQVTVGSPSSP-----LAVQPMANVINSANGFGAICOLNV 1243  
 QY 1244 YNNKASGSSRRRRRSIQNOEAPDLVAVKENKODLNHVDLNVCTSPSGRSGMALMEVNL 1303  
 DB 1244 YNNKASGSSRRRRRSIQNOEAPDLVAVKENKODLNHVDLNVCTSPSGRSGMALMEVNL 1303  
 QY 1304 LSGFVAVSEALISLSTYKVEYDQKLNLYLDSVNETOPCVNIPAVRPFKUSNTODASVS 1363  
 DB 1304 LSGFVAVSEALISLSTYKVEYDQKLNLYLDSVNETOPCVNIPAVRPFKUSNTODASVS 1363  
 QY 1364 IVDYIEPRROAVRSYNSSEVKLSCDLCSDVQCGRCCEBGASGSHHSVYIFFCFKLYF 1423  
 DB 1364 IVDYIEPRROAVRSYNSSEVKLSCDLCSDVQCGRCCEBGASGSHHSVYIFFCFKLYF 1423  
 QY 1424 MELWL 1428  
 DB 1441 MELWL 1445

RESULT 3  
ID ABB82167 standard; Protein: 1445 AA.  
XX ABB82167;  
AC ABB82167;  
XX 23-DEC-2002 (first entry)  
DT 23-DEC-2002 (first entry)  
DE Human CD109 KI-H7 protein.  
XX  
XX CD109, alpha2 macroglobulin; thioester; cerebroprotective; cardiant;  
KM immunosuppressive; haemostatic; anticoagulant; thrombolytic; human;  
KM cardiovascular; vasotropic; gene therapy; CD109 KI-H7.  
OS Homo sapiens.  
XX  
XX WO200270696-A2.  
XX  
XX 12-SEP-2002.  
XX  
XX 07-MAR-2002; 2002MO-CA00292.  
XX  
XX 07-MAR-2001; 2001US-273814P.  
XX  
XX (SCHU/) SCHUH A.  
PA (SUTH/) SUTHERLAND R D.  
XX  
XX Schuh A, Sutherland RD;  
XX  
XX WPI; 2002-713450/77.  
DR N-PSDB; ABQ79966.  
XX  
XX New CD109 nucleic acid and polypeptides, useful in gene therapy,  
PT particularly for treating strokes, myocardial infarctions, thrombosis,  
PT thrombocytopenia, autoimmune diseases, or organ or bone marrow  
PT transplantation -  
XX  
XX Claim 21; Fig 3a; 156pp; English.  
XX  
XX The invention relates to isolated nucleic acid molecules encoding CD109  
XX polypeptides. These nucleic acid molecules include the human cDNA  
XX sequences comprising CD109 KI, CD109 KI-H7, CD109 K15 or their variants.  
XX CD109 is a novel member of the alpha2 macroglobulin (alpha2M)/C3, C4, C5  
XX family of thioester-containing proteins. The CD109 polypeptides can be  
XX expressed by standard recombinant methodology. The CD109 nucleic acid,  
XX CD109 polypeptide or its mimetic, or the anti-CD109 antibody is useful  
XX for treating or detecting a disease or disorder, e.g. conditions  
XX associated with endothelial activation, platelet activation, activation  
XX of the coagulation or fibrinolytic systems, activation of T lymphocytes  
XX and of the complement system, quantitative or qualitative abnormalities  
XX of platelet function, increased or impaired platelet aggregation and  
XX activation, increased or impaired activation of the coagulation and/or  
XX fibrinolytic systems, or impaired or increased immune activation. These  
XX are also useful for treating cardiovascular disorders, stroke, myocardial  
XX infarction, thrombosis, embolism, peripheral vascular disease,  
XX thrombocytopenia, thrombocytopenia, autoimmune diseases, organ  
XX transplantation, or bone marrow transplantation. The CD109 nucleic acids  
XX or their antisense nucleotide sequence are useful in gene therapy for  
XX treating these conditions. The present sequence represents the human  
XX CD109 KI-H7 sequence.  
XX  
XX Sequence 1445 AA;  
SQ

Query Match 99.7%; Score 7326.5; DB 23; Length 1445;  
Best Local Similarity 98.8%; Pred. No. 0;  
Matches 1427; Conservative 1; Mismatches 0; Indels 17; Gaps 1;

QY 1 MOSPRLTAHLTCVCAALAVAPGRFLVATPRTIPGNVNTGVLLHSCSQVTYKA 60  
DB 1 MOSPRLTAHLTCVCAALAVAPGRFLVATPRTIPGNVNTGVLLHSCSQVTYKA 60

QY 61 ELIKTASNLTVSLVLAEGVPEKGSFKTTLPLPSLPLNSADEIYELRYTGTODEILPSNST 120  
DB 61 ELIKTASNLTVSLVLAEGVPEKGSFKTTLPLPSLPLNSADEIYELRYTGTODEILPSNST 120  
QY 121 RLSFETKRISVFTQTKALYKPKQEVKFRIVLTFSPFKPKTSINTLIDPKSNLQQWL 180  
DB 121 RLSFETKRISVFTQTKALYKPKQEVKFRIVLTFSPFKPKTSINTLIDPKSNLQQWL 180  
QY 181 SQQSDLGIVISKTPQLSSHPILGDMISIQVNDQTYQSFQVSEYVLPKEVTLQTLVYCS 240  
DB 181 SQQSDLGIVISKTPQLSSHPILGDMISIQVNDQTYQSFQVSEYVLPKEVTLQTLVYCS 240  
QY 241 MNSKHLNGITITAKTYGKPKVGDVTLPLPLSFWKKKKNTTKFKINGSANFSFNDENK 300  
DB 241 MNSKHLNGITITAKTYGKPKVGDVTLPLPLSFWKKKKNTTKFKINGSANFSFNDENK 300  
QY 301 NMDSSNGLSSEYLDSSPGVEILTITVTSVNGISNNVSTNVEFKQHDYIEFDDYTVL 360  
DB 301 NMDSSNGLSSEYLDSSPGVEILTITVTSVNGISNNVSTNVEFKQHDYIEFDDYTVL 360  
QY 361 KPSLNTFAITVKTTRADGNQTLTEERRNNVITYTQNTYEYWSGNSGNQKMAVOKINY 420  
DB 361 KPSLNTFAITVKTTRADGNQTLTEERRNNVITYTQNTYEYWSGNSGNQKMAVOKINY 420  
QY 421 TVPQSGTFKIEPILDSSELQKAYFLGSKSMVAHSLFKSPSKTYIQLKTRDENIKVG 480  
DB 421 TVPQSGTFKIEPILDSSELQKAYFLGSKSMVAHSLFKSPSKTYIQLKTRDENIKVG 480  
QY 481 SPFELVSGNKRKLELSYMWVSRGQLVAVGKONSTFSLTPENSWPKACVITYYIEDDG 540  
DB 481 SPFELVSGNKRKLELSYMWVSRGQLVAVGKONSTFSLTPENSWPKACVITYYIEDDG 540  
QY 541 EISIDVLKIPVOLVFNKIKILYMSKYKAESEKVSIRISVTOPDSIVGIVAVDKSVNLNM 600  
DB 541 EISIDVLKIPVOLVFNKIKILYMSKYKAESEKVSIRISVTOPDSIVGIVAVDKSVNLNM 600  
QY 601 ASNDITMENVHELELYNTGYIYGMFNSPAVQECGLWVLDANITKQYIDGVYNAEY 660  
DB 601 ASNDITMENVHELELYNTGYIYGMFNSPAVQECGLWVLDANITKQYIDGVYNAEY 660  
QY 661 AERFMEENGHIYDIDHPSLGSPPHYRKHPEPWIMLDNMGRIYQEPVAVVPDSITSW 720  
DB 661 AERFMEENGHIYDIDHPSLGSPPHYRKHPEPWIMLDNMGRIYQEPVAVVPDSITSW 720  
QY 721 VARGFISEDLGLGTLTTPVELOAFQPFILNLPSVIRGEEFALEITFYNLKXATEY 780  
DB 721 VARGFISEDLGLGTLTTPVELOAFQPFILNLPSVIRGEEFALEITFYNLKXATEY 780  
QY 781 KVIEKSDKFDILMTSSEINATGHOQTLVPSEDAVLPPIRPHLGEIPITVTLASPT 840  
DB 781 KVIEKSDKFDILMTSSEINATGHOQTLVPSEDAVLPPIRPHLGEIPITVTLASPT 840  
QY 841 ASDAVTOMILVRAEGIEKYSQSILDLTDNRLOSTLKTLSRFPNTYGSRRVITAI 900  
DB 841 ASDAVTOMILVRAEGIEKYSQSILDLTDNRLOSTLKTLSRFPNTYGSRRVITAI 900  
QY 901 GDTLGSINGLASLIMPGCCGQNNINFAPNITIDYTKKKQLDNLKEXALSMRQO 960  
DB 901 GDTLGSINGLASLIMPGCCGQNNINFAPNITIDYTKKKQLDNLKEXALSMRQO 960  
QY 961 YORELLYOREDEGSFSAFGNYDPDSGSTMLSAFVLRCELEADPYIDIDQNVLAHRTYTLKGH 1020  
DB 961 YORELLYOREDEGSFSAFGNYDPDSGSTMLSAFVLRCELEADPYIDIDQNVLAHRTYTLKGH 1020  
QY 1021 QKNGSEFMDPGRVYHSELQGNKSPVTLTAIYVTSILGRTKQPNIDVQESHIFLESSEFS 1080  
DB 1021 QKNGSEFMDPGRVYHSELQGNKSPVTLTAIYVTSILGRTKQPNIDVQESHIFLESSEFS 1080  
QY 1081 RGISDNYTLALITYALSSVSPKAKELNMLTMRAQEGGOMQWVSESGLSDMQPRSL 1140  
DB 1081 RGISDNYTLALITYALSSVSPKAKELNMLTMRAQEGGOMQWVSESGLSDMQPRSL 1140  
QY 1141 DIEVAAYALLSHFLQFOTSEGPIPMRWLSRORNSLGSFASODDTVALKALSEFALMNT 1200

```

Db      1141 DIEVAAYALSHFLQFOTISEGIPIMRWLSRORNSJGGFASDTDDTVALKAUSEFALMNT 1200
Qy      1201 ERTNIQVTVTPSSBSP-----LAVVQPMANISANGFGAICQLNV 1243
Db      1201 ERTNIQVTVTPSSBSPVKFLIDITNRLLQTAELAVVQPMANISANGFGAICQLNV 1260
Qy      1244 YNVKASGSSRRRRRSIQNOEAPFLDVAVKENKODLNHVDLNVCTSFSGPGRSGMALMEVNL 1303
Db      1261 YNVKASGSSRRRRRSIQNOEAPFLDVAVKENKODLNHVDLNVCTSFSGPGRSGMALMEVNL 1320
Qy      1304 LSGFNVPEEATSLSTVWKVVEVDHCKLNLXLDNVNETQFCVNIIPAVRNPFKVSNTODASYS 1363
Db      1321 LSGFNVPEEATSLSTVWKVVEVDHCKLNLXLDNVNETQFCVNIIPAVRNPFKVSNTODASYS 1380
Qy      1364 IVDYVEPRROAVRSYNSEVKLSGCDLCSGVQGCRCPCEDASGSHHSVITIFCFKLLYF 1423
Db      1381 IVDYVEPRROAVRSYNSEVKLSGCDLCSGVQGCRCPCEDASGSHHSVITIFCFKLLYF 1440
Qy      1424 MELWL 1428
Db      1441 MELWL 1445

```

## RESULT 4

AA019372 ID AA019372. standard; Protein; 1445 AA.

AA019372;

04-DEC-2002 (first entry)

Human platelet alloantigen Gova.

Human; platelet alloantigen; Gova; Govb; single nucleotide polymorphism; SNP; biallelic; bleeding disorder; post-transfusion purpura; post-transfusion platelet refractoriness; haemostatic; vaccine; neonatal alloimmune thrombocytopenia.

Homo sapiens.

MO200270738-A2.

12-SEP-2002.

07-MAR-2002; 2002MO-CA00291.

07-MAR-2001; 2001US-273941P.

(SCHU// SCHUH A.

Schuh A, Ouehand W;

WPI; 2002-713460/77.

N-PSDB; AAL49815.

New isolated oligonucleotide binding to a region of CD109 nucleic acid having a single nucleotide polymorphism that distinguishes a Gova and/or Govb allele, useful for treating blood disorders e.g. alloimmune thrombocytopenia

Disclosure; Page 29-35; 69pp; English.

The present invention relates to a sequence capable of binding specifically to a CD109 nucleic acid which has a single nucleotide polymorphism that distinguishes the Gova and Govb alleles. Detection of the Gov genotype is useful for detecting whether the subject has or is at risk of a blood disease, disorder or abnormal physical state, such as bleeding, or increased risk of bleeding, due to alloimmune destruction of blood platelets, e.g., post-transfusion purpura, post-transfusion platelet refractoriness or neonatal alloimmune thrombocytopenia. The nucleic acid and polypeptide are useful for Gov genotyping or phenotyping individuals. The present sequence is the Gova encoded protein.

XX Sequence 1445 AA;  
SQ  
Query Match 99.7%; Score 7326.5; DB 23; Length 1445;  
Best Local Similarity 98.8%; Pred. No. 0;  
Matches 1427; Conservative 1; Mismatches 0; Indels 17; Gaps 1;

```

Qy      1 MOGPBLTAAHLCTCTAALAVAPGPRELVTAAGIIRPGNVITGVLELHCPGQVTKA 60
Db      1 MOGPBLTAAHLCTCTAALAVAPGPRELVTAAGIIRPGNVITGVLELHCPGQVTKA 60
Qy      61 ELTKASNLTVSVLEAEVFEKSGFKITLPQLPNSADEIYEILRVGTODEILFNSNT 120
Db      61 ELTKASNLTVSVLEAEVFEKSGFKITLPQLPNSADEIYEILRVGTODEILFNSNT 120
Qy      121 RLSFETKRSIVFIQTDKALYRKQEVKRYITLPSDPFPYTSINILKDKPSNLIQWL 180
Db      121 RLSFETKRSIVFIQTDKALYRKQEVKRYITLPSDPFPYTSINILKDKPSNLIQWL 180
Qy      181 SQQSDLAGVISKTFQLSHPILGWSIQVQVNDQTYGQFQSEVYLPRFEYTLQTPLYCS 240
Db      181 SQQSDLAGVISKTFQLSHPILGWSIQVQVNDQTYGQFQSEVYLPRFEYTLQTPLYCS 240
Qy      241 MNSKRLNGTITAKTYGKPKVGDVTLTFLPLSPFKKKNIITKPKINGSANFSFDEEMK 300
Db      241 MNSKRLNGTITAKTYGKPKVGDVTLTFLPLSPFKKKNIITKPKINGSANFSFDEEMK 300
Qy      301 NVMDSSNGLSFYLDSFGPVEILTTVYESVTGISRNVSTVFRKHQYIIIEFDYTVL 360
Db      301 NVMDSSNGLSFYLDSFGPVEILTTVYESVTGISRNVSTVFRKHQYIIIEFDYTVL 360
Qy      361 KPSLNFATVATVTRADGNQLTLERRNNVITVQRNTEYWSGNSGNQKMEAVOKINY 420
Db      361 KPSLNFATVATVTRADGNQLTLERRNNVITVQRNTEYWSGNSGNQKMEAVOKINY 420
Qy      421 TVPQSGTFKIEPPILEDSSSELQKAYFLGSKSMAVHSLFKSPKXTYIQLKRDENIKYG 480
Db      421 TVPQSGTFKIEPPILEDSSSELQKAYFLGSKSMAVHSLFKSPKXTYIQLKRDENIKYG 480
Qy      481 SPFELVYSGNRKLELSYMYVSRQQLVAVGQNSTMTSELTPENSGTTPACVIVYIEEDG 540
Db      481 SPFELVYSGNRKLELSYMYVSRQQLVAVGQNSTMTSELTPENSGTTPACVIVYIEEDG 540
Qy      541 EIIISDLKIPVQLYPKNKIKLYMSKVKAPSEKYSLSRSTVQPSIVGIVAVDSVNLNN 600
Db      541 EIIISDLKIPVQLYPKNKIKLYMSKVKAPSEKYSLSRSTVQPSIVGIVAVDSVNLNN 600
Qy      601 ASNDITMENVHELELYNTGYLLGMFNNSFAVFOEGLMVLTDLNLTQDYIDGYVDNAEY 660
Db      601 ASNDITMENVHELELYNTGYLLGMFNNSFAVFOEGLMVLTDLNLTQDYIDGYVDNAEY 660
Qy      661 AERFMEENEGHIVDIHDFSLGSSPHVRKHPERTIMWLDTNNGRIYQEFVTVDPDSITSW 720
Db      661 AERFMEENEGHIVDIHDFSLGSSPHVRKHPERTIMWLDTNNGRIYQEFVTVDPDSITSW 720
Qy      721 VATGFVISEDGLGTTTPVELQAFOPPEFLNIPYSIRKEEPLAETITNNYLKDADEV 780
Db      721 VATGFVISEDGLGTTTPVELQAFOPPEFLNIPYSIRKEEPLAETITNNYLKDADEV 780
Qy      781 KVILEKSDKFDILMTSSSEINATGHOQTLVSESDGATVLPFIRPTHGIEIITVYALSPT 840
Db      781 KVILEKSDKFDILMTSSSEINATGHOQTLVSESDGATVLPFIRPTHGIEIITVYALSPT 840
Qy      841 ASDAVTQMLIVKAGIEKYSQSIILDLITDNRLOSTKTSFSPPNVTGSSERVQITAI 900
Db      841 ASDAVTQMLIVKAGIEKYSQSIILDLITDNRLOSTKTSFSPPNVTGSSERVQITAI 900
Qy      901 GDVIGPSINGLASLIRMPYGGGEOMTNAFANIYILVLTCKKQLTDLKKAISPFMOG 960
Db      901 GDVIGPSINGLASLIRMPYGGGEOMTNAFANIYILVLTCKKQLTDLKKAISPFMOG 960
Qy      961 YORELLYOREDDGSAFAGNDYPSGTWLSAFLRCFLADPYIDIDQNVLRRTYTWLKGH 1020
Db      961 YORELLYOREDDGSAFAGNDYPSGTWLSAFLRCFLADPYIDIDQNVLRRTYTWLKGH 1020

```

Db 961 YQRELLYQREDEGSFSAFGNYPDSGTWLSAFLVLCFLEADPYDIDQNVLHRTYTWLKGH 1020  
Qy 1021 QKSGEFPMDGRVHSELQGGNKSPTLTAYIVTSLGYRKYQPNIDVQESIHFLSESEFS 1080  
Db 1021 QKSGEFPMDGRVHSELQGGNKSPTLTAYIVTSLGYRKYQPNIDVQESIHFLSESEFS 1080  
Qy 1081 RGISDNTTALTALTYALS SVGS PKAKELNMLTWRABEGQMGQFVWSESSEKLSDSWQPSRL 1140  
Db 1081 RGISDNTTALTALTYALS SVGS PKAKELNMLTWRABEGQMGQFVWSESSEKLSDSWQPSRL 1140  
Qy 1141 DIEVAALSHFLQOPTSGSIPIMRWLSHQRNLSGFPASTQDTTVALKALSEFALMNT 1200  
Db 1141 DIEVAALSHFLQOPTSGSIPIMRWLSHQRNLSGFPASTQDTTVALKALSEFALMNT 1200  
Qy 1201 ERTNIQVTVTGSPSSPSP-----LAVQPMVANI SANGCFALCQLNVV 1243  
Db 1201 ERTNIQVTVTGSPSSPSP-----LAVQPMVANI SANGCFALCQLNVV 1243  
Qy 1244 YNVKASGSSRRRRSIQNOEAFDLDAVAKENKDNLDVAVNCTSSFGSPGSGMALMEVNL 1303  
Db 1261 YNVKASGSSRRRRSIQNOEAFDLDAVAKENKDNLDVAVNCTSSFGSPGSGMALMEVNL 1320  
Qy 1304 LSGFMVSESAISLSEYTKYKVEYDHGKNTLYDSVNEQFCVNI PAVNPFKVSNTQDASVS 1363  
Db 1321 LSGFMVSESAISLSEYTKYKVEYDHGKNTLYDSVNEQFCVNI PAVNPFKVSNTQDASVS 1380  
Qy 1364 IVDYEBRQAVRNSYSEVKLSGDLCSDPVQGRPCEDGASGSHHSSVIFIFCFKLLYF 1423  
Db 1381 IVDYEBRQAVRNSYSEVKLSGDLCSDPVQGRPCEDGASGSHHSSVIFIFCFKLLYF 1440  
Qy 1424 MELWL 1428  
Db 1441 MELWL 1445

## RESULT 5

AAE32013 standard; Protein; 1445 AA.

ID AAE32013;  
AC AAE32013;DT 24-MAR-2003 (first entry)  
DE Human r150 protein #2.KW Human; GPI-anchored TGF-beta1 binding protein; r150 protein; cancer;  
glycosylphosphatidyl inositol; transforming growth factor-beta1;  
therapy.  
XX Homo sapiens.  
OS  
FH Key Location/Qualifiers  
FT Misc-difference 703 /label= Ser, Tyr  
FT /note= "Encoded by TNC"XX PN W0200285942-A2.  
XX PD 31-OCT-2002.  
XX PF 24-APR-2002; 2002WO-CA00560.  
XX PR 24-APR-2001; 2001US-285713P.  
XX PR 14-FEB-2002; 2002US-356163P.XX (UYMC-) UNIV MCGIL.  
XX PI Philip A, Tam B;  
XX WPI; 2003-093100/08.  
XX DR N-PSDB; AAD49435.  
XX PT Novel transforming growth factor (TGF)-beta 1 binding reagent which

PT comprises r150 protein which acts as accessory receptor of TGF-beta,  
PT useful for negatively modulating TGF-beta activity, and thus for  
PT treating cancer -  
PS Claim 1; Page 103-109; 127pp; English.  
XX The invention relates to novel transforming growth factor (TGF)-beta1  
CC binding reagent which comprises a glycosylphosphatidyl inositol (GPI)-  
CC anchored TGF-beta1 binding protein referred to as r150 which acts as  
CC an accessory receptor of TGF-beta. The invention is used for negatively  
CC modulating TGF-beta activity, and thus for treating conditions  
CC characterised by overproduction of TGF-beta, such as cancer. Antisense  
CC molecules of the invention are used for increasing TGF-beta availability  
CC and increase graft success. The present sequence is human r150 protein.  
XX  
SQ Sequence 1445 AA;

Query Match 99.6%; Score 7318.5; DB 24; Length 1445;

Best Local Similarity 98.7%; Pred. No. 0;

Matches 1426; Conservative 1; Mismatches 1; Indels 17; Gaps 1;

Qy 1 MOGPPLTAAHLCTCTAALAVAGPRPLVTARGIIRPGSNVTIGVLEHCHSOYTVKA 60  
Db 1 MOGPPLTAAHLCTCTAALAVAGPRPLVTARGIIRPGSNVTIGVLEHCHSOYTVKA 60  
Qy 61 ELIKTASNLVSLAEAGVPEKESFKTLPLSLPLNSADEIYELARTGRTOEILFSNST 120  
Db 61 ELIKTASNLVSLAEAGVPEKESFKTLPLSLPLNSADEIYELARTGRTOEILFSNST 120  
Qy 121 RLSFETKRISVFIQTDKALYKQEVKFRIVTLFSPDKPKYKSLNLIKOPKSNLIQOWL 180  
Db 121 RLSFETKRISVFIQTDKALYKQEVKFRIVTLFSPDKPKYKSLNLIKOPKSNLIQOWL 180  
Qy 181 SQOSDLGVSKTEFQLSHPILGDMSTIOVQNDQTYQSFQVSEYVLPKPEVTLQTPLYCS 240  
Db 181 SQOSDLGVSKTEFQLSHPILGDMSTIOVQNDQTYQSFQVSEYVLPKPEVTLQTPLYCS 240  
Qy 241 MNSKHLNGITTAKYTGKPYKGDVTLTFLPLSPFGKKKNTTKFKINGSANFSFDEMK 300  
Db 241 MNSKHLNGITTAKYTGKPYKGDVTLTFLPLSPFGKKKNTTKFKINGSANFSFDEMK 300  
Qy 301 NWMDSSNGLSSEYDLSSPGVEILTTVTESVTGISRNVSTNVFPKHDIIEFDDYTVL 360  
Db 301 NWMDSSNGLSSEYDLSSPGVEILTTVTESVTGISRNVSTNVFPKHDIIEFDDYTVL 360  
Qy 361 KPSLNTATATKYTRAGNOLTLERRNNVITYTORNYEYWGSGNSGNQKMAVQKINY 420  
Db 361 KPSLNTATATKYTRAGNOLTLERRNNVITYTORNYEYWGSGNSGNQKMAVQKINY 420  
Qy 421 TVPQSGTFKIEFPILEDSELOKAYFLGSKSSMAVHSLFKSPSKTYIQKTRDENIKVG 480  
Db 421 TVPQSGTFKIEFPILEDSELOKAYFLGSKSSMAVHSLFKSPSKTYIQKTRDENIKVG 480  
Qy 481 SPFELVSGNKRKLEISYVWSRGQLVAVGKQNSTWFSLTPENSWPKACVIVYIEDDG 540  
Db 481 SPFELVSGNKRKLEISYVWSRGQLVAVGKQNSTWFSLTPENSWPKACVIVYIEDDG 540  
Qy 541 EIIISDVLYKIPVQVFPKKIKLYWSKYKABSEKVSIRISTQPDSTVGIYAVDKSVNLM 600  
Db 541 EIIISDVLYKIPVQVFPKKIKLYWSKYKABSEKVSIRISTQPDSTVGIYAVDKSVNLM 600  
Qy 601 ASNDITMENVHELELYNTGYLYGMFMSFAVFOEGCLMWLTDAKITKIDYDYDAAEY 660  
Db 601 ASNDITMENVHELELYNTGYLYGMFMSFAVFOEGCLMWLTDAKITKIDYDYDAAEY 660  
Qy 661 AERFMEENEGHIVDHFSLGSSPHVKRHPETWIMLDTWNGYRIYQEFVTVPDSITSW 720  
Db 661 AERFMEENEGHIVDHFSLGSSPHVKRHPETWIMLDTWNGYRIYQEFVTVPDSITSW 720  
Qy 721 VATGFVISEDGLGLTTTPVELQAFQPFIFLNLPSVINGEFALEITTFNLIKQATEV 780  
Db 721 VATGFVISEDGLGLTTTPVELQAFQPFIFLNLPSVINGEFALEITTFNLIKQATEV 780



```

QY 781 KVIIEKSDKFDILMTSSEINATCHOQTLLVPSFDGATVLPPIRPTHGLSIPITVTALSTP 840
DB 781 KVIIEKSDKFDILMTSSEINATCHOQTLLVPSFDGATVLPPIRPTHGLSIPITVTALSTP 840
QY 841 ASDAVTOMILVAEGIEKSYSOSILLDLIDNRLQSTLTKLSPSPNTWGSERVOTIT 900
DB 841 ASDAVTOMILVAEGIEKSYSOSILLDLIDNRLQSTLTKLSPSPNTWGSERVOTIT 900
QY 901 GDLVPSINGLASLIRMPYCGEQQNNINPAPNIYIIDLYLTKKKQLTDNLKEKALSFMRQ 960
DB 901 GDLVPSINGLASLIRMPYCGEQQNNINPAPNIYIIDLYLTKKKQLTDNLKEKALSFMRQ 960
QY 961 YORELLYQREDDSPFAFGNYDPSGSTMALFVLRCLFADPYIIDQNLHRTYTWLKH 1020
DB 961 YORELLYQREDDSPFAFGNYDPSGSTMALFVLRCLFADPYIIDQNLHRTYTWLKH 1020
QY 1021 QKSNGEFMDPGRVHSELQGNKSPVTLTAYIVTSLIGRKYQNPINDQESIHFESEFS 1080
DB 1021 QKSNGEFMDPGRVHSELQGNKSPVTLTAYIVTSLIGRKYQNPINDQESIHFESEFS 1080
QY 1081 RGISDNYTLALITVALSYSGSPKAKALNMLTWRAEOEGMOPWVSSEKLSDSMQPRSL 1140
DB 1081 RGISDNYTLALITVALSYSGSPKAKALNMLTWRAEOEGMOPWVSSEKLSDSMQPRSL 1140
QY 1141 DIEVAAYALLSHFLOFQISEGIPIMKWLSRQNSLGGFASDTDTVALKALSEFALNMT 1200
DB 1141 DIEVAAYALLSHFLOFQISEGIPIMKWLSRQNSLGGFASDTDTVALKALSEFALNMT 1200
QY 1201 ERTNIQVYTYGSPSSPSP-----LAVQPMANISANGFPAICQLNV 1243
DB 1201 ERTNIQVYTYGSPSSPSPKFLIDTNHRLLLQTAELAVQPMANISANGFPAICQLNV 1260
QY 1244 YVVKASGSSRRRRSIIQNOAFDLDAVAKENKDLNHDVNLNCTSFSGPGRSGMALMEVNL 1303
DB 1244 YVVKASGSSRRRRSIIQNOAFDLDAVAKENKDLNHDVNLNCTSFSGPGRSGMALMEVNL 1303
QY 1304 LSGFVWPSEAIISLSETVKVEYDHGKLANLYLDSVNETQFCVNI PAVRNKVSNTODASVS 1363
DB 1304 LSGFVWPSEAIISLSETVKVEYDHGKLANLYLDSVNETQFCVNI PAVRNKVSNTODASVS 1380
QY 1364 IVDYIEPRRQAVRSYNSVYKLSGCDLCSVQCGRPEDDASGSHHSSVITFICFGLIYF 1423
DB 1364 IVDYIEPRRQAVRSYNSVYKLSGCDLCSVQCGRPEDDASGSHHSSVITFICFGLIYF 1440
QY 1424 MELML 1428
DB 1441 MELML 1445

```

```

PA (SCHU/) SCHUH A.
PA (SUTH/) SUTHERLAND R D.
PI Schuh A, Sutherland RD;
XX MPI, 2002-713450/77.
DR N-PSDB; ABQ79965.
XX
XX New CD109 nucleic acids and polypeptides, useful in gene therapy,
PT particularly for treating strokes, myocardial infarctions, thrombosis,
PT thrombocytopenia, autoimmune diseases, or organ or bone marrow
PT transplantation -
XX
XX Claim 21; Fig 3b; 156bp; English.
XX
CC The invention relates to isolated nucleic acid molecules encoding CD109
CC polypeptides. These nucleic acid molecules include the human cDNA
CC sequences comprising CD109 KI, CD109 KI-H7, CD109 K15 or their variants.
CC CD109 is a novel member of the alpha2 macroglobulin (alpha2M)/C3, C4, C5
CC family of thioester-containing proteins. The CD109 polypeptides can be
CC expressed by standard recombinant methodology. The CD109 nucleic acid,
CC CD109 polypeptide or its mimetic, or the anti-CD109 antibody is useful
CC for treating or detecting a disease or disorder, e.g. conditions
CC associated with endothelial activation, platelet activation, activation
CC of the coagulation or fibrinolytic systems, activation of T lymphocytes
CC and of the complement system, quantitative or qualitative abnormalities
CC of platelet function, increased or impaired platelet aggregation and/or
CC activation, increased or impaired activation of the coagulation and/or
CC fibrinolytic systems, or impaired or increased immune activation. These
CC are also useful for treating cardiovascular disorders, stroke, myocardial
CC infarction, thrombosis, embolism, peripheral vascular disease,
CC thrombocytopenia, thrombocythemia, autoimmune diseases, organ
CC transplantation, or bone marrow transplantation. The CD109 nucleic acids
CC or their antisense nucleotide sequence are useful in gene therapy for
CC treating these conditions. The present sequence represents the human
CC CD109 KI variant sequence.
XX
SQ Sequence 1445 AA;

```

Query Match 99.6%; Score 7317.5; DB 23; Length 1445;  
 Best Local Similarity 98.7%; Pred. No. 0;  
 Matches 1426; Conservative 1; Mismatches 1; Indels 17; Gaps 1;

```

QY 1 MCGPPLTAAHLCTCTAALAAVAPGRPLVTPAGIIRPGANTVIGELLEHSPQVTVAA 60
DB 1 MCGPPLTAAHLCTCTAALAAVAPGRPLVTPAGIIRPGANTVIGELLEHSPQVTVAA 60
QY 61 ELKKTASNLTVSVLEAEVFEKSGFKTLTPSLPLNSADEIYELRVYGRTODEILFSNST 120
DB 61 ELKKTASNLTVSVLEAEVFEKSGFKTLTPSLPLNSADEIYELRVYGRTODEILFSNST 120
QY 121 RLSFETKRISVFIQTDKALKYKQVKKRIVTLBSDFPKYKTSNLILIDPKPSNLIQQL 180
DB 121 RLSFETKRISVFIQTDKALKYKQVKKRIVTLBSDFPKYKTSNLILIDPKPSNLIQQL 180
QY 181 SQQSDLVGIVSKTFOLSSHPILGDSIIOVANDQTYGOSFOVEEYLPKREYVLTQTPLYS 240
DB 181 SQQSDLVGIVSKTFOLSSHPILGDSIIOVANDQTYGOSFOVEEYLPKREYVLTQTPLYS 240
QY 241 MNSKHLNGITIAKTYTGKPVKDVLTLPPLSPFKKKKNIITFKINGSANSPFNDEMK 300
DB 241 MNSKHLNGITIAKTYTGKPVKDVLTLPPLSPFKKKKNIITFKINGSANSPFNDEMK 300
QY 301 NMDSNGISEYLDSSPEVEILTTVBSVTGISRANSTNVFFQOHDYIIIEFDYTVL 360
DB 301 NMDSNGISEYLDSSPEVEILTTVBSVTGISRANSTNVFFQOHDYIIIEFDYTVL 360
QY 361 KPSLNFATVAVKTRADGNQLTLEERRNNVLTVOQNTVEWGSNGSKQKEAVQKINY 420
DB 361 KPSLNFATVAVKTRADGNQLTLEERRNNVLTVOQNTVEWGSNGSKQKEAVQKINY 420
QY 421 TVPQSGFKIEPILDESSSELQKAYPLGSKSMAVHSLFKSPSKTYIQLKTRDENIKVG 480
DB 421 TVPQSGFKIEPILDESSSELQKAYPLGSKSMAVHSLFKSPSKTYIQLKTRDENIKVG 480

```

```

Db 421 TVPQSGTFKIEPFLIEDSSSELQLAAYFLGSKSMVAHSLFKPSKTYIQLKTRDENIKVG 480
Qy 481 SPFELVVSQNKRLKELSYMVVSRGOLVAVGKONSTWESLTPENSWTPKACIYVYIEDDG 540
Db 481 SPFELVVSQNKRLKELSYMVVSRGOLVAVGKONSTWESLTPENSWTPKACIYVYIEDDG 540
Qy 541 EIIISDVLIKIPVQVLPKNNKIKLYMSKVYAEPESEKVSRLISVTPDSIVGIVAVDKSVLNM 600
Db 541 EIIISDVLIKIPVQVLPKNNKIKLYMSKVYAEPESEKVSRLISVTPDSIVGIVAVDKSVLNM 600
Qy 601 ASNIITMENVYHELVNTGYGGLGFMFMSPAVPOECGLMUTDANLTKDYIDGYDYNAEX 660
Db 601 ASNDITMENVYHELVNTGYGGLGFMFMSFAVPOECGLMUTDANLTKDYIDGYDYNAEX 660
Qy 661 AERFMEENEGHIVIHDPFSLGSSPHVRKHFPETIMIDTNNGYRIYOEFEVTVDSITSW 720
Db 661 AERFMEENEGHIVIHDPFSLGSSPHVRKHFPETIMIDTNNGYRIYOEFEVTVDSITSW 720
Qy 721 VATGFVISBDLGLITTPVELQAPQPFIFLNLPSYVIRGEFPALEITTFNVLKDATEV 780
Db 721 VATGFVISBDLGLITTPVELQAPQPFIFLNLPSYVIRGEFPALEITTFNVLKDATEV 780
Qy 781 KVIIEKSDKPDIIIMTSEFINATGHOQTLVPSDEGATVLPFIRPTHIGEIPITVTAISPT 840
Db 781 KVIIEKSDKPDIIIMTSEFINATGHOQTLVPSDEGATVLPFIRPTHIGEIPITVTAISPT 840
Qy 841 ASDAVTQMLVKAEGIEKYSQSILDLTDNRLOSTLKTSSFPSPPTVTGSEKVOITAI 900
Db 841 ASDAVTQMLVKAEGIEKYSQSILDLTDNRLOSTLKTSSFPSPPTVTGSEKVOITAI 900
Qy 901 GDVLGPSINGLASLIRMPYCGEONMINFAPNIYILDYLTKKQOLTNLKEKALSPKROG 960
Db 901 GDVLGPSINGLASLIRMPYCGEONMINFAPNIYILDYLTKKQOLTNLKEKALSPKROG 960
Qy 961 YQRELLIYQREBGSASAGNDPSSGTLASAVLRCFLPADYIDIDQNVLHRTYTWLKGH 1020
Db 961 YQRELLIYQREBGSASAGNDPSSGTLASAVLRCFLPADYIDIDQNVLHRTYTWLKGH 1020
Qy 1021 OKSGEFPDGRVHSELOGNKSPTVLTAYITVSLGIRKYONIDVOESIHFLSESPS 1080
Db 1021 OKSGEFPDGRVHSELOGNKSPTVLTAYITVSLGIRKYONIDVOESIHFLSESPS 1080
Qy 1081 RGISDNTYLLALITYALS SVGS PKAKKALNMLTWRAEBOGQMFVSSBSKLSDSMOPRSL 1140
Db 1081 RGISDNTYLLALITYALS SVGS PKAKKALNMLTWRAEBOGQMFVSSBSKLSDSMOPRSL 1140
Qy 1141 DIEVAAYALLSHFQFOTSEGIPIRMWLSRORNSLGGFASQDPTVALKALSEFPALMNT 1200
Db 1141 DIEVAAYALLSHFQFOTSEGIPIRMWLSRORNSLGGFASQDPTVALKALSEFPALMNT 1200
Qy 1201 ERTNIQVTVGSPSSPSP-----LAVVQPAVAVNISANGFGFALCOLNVV 1243
Db 1201 ERTNIQVTVGSPSSPSPKFLIDITHNRLLOTAEILAVVQPAVAVNISANGFGFALCOLNVV 1243
Qy 1244 YNVKASSSRRRRSIQNOEAFDLDAVAKENDDLNHDVNLVCTSFSGRSGMALMEVNL 1303
Db 1244 YNVKASSSRRRRSIQNOEAFDLDAVAKENDDLNHDVNLVCTSFSGRSGMALMEVNL 1303
Qy 1261 YNVKASSSRRRRSIQNOEAFDLDAVAKENDDLNHDVNLVCTSFSGRSGMALMEVNL 1320
Db 1261 YNVKASSSRRRRSIQNOEAFDLDAVAKENDDLNHDVNLVCTSFSGRSGMALMEVNL 1320
Qy 1304 LSGFMVSEALISLSEYVAKVEYDHGKLNLYLDSVNETQFCNIAVAVNFKVSNQDASVS 1363
Db 1304 LSGFMVSEALISLSEYVAKVEYDHGKLNLYLDSVNETQFCNIAVAVNFKVSNQDASVS 1363
Qy 1364 IVDYEEPRROAVRSYNSRVKLSGDLCSVOGCRPCEDGASGSHHSVIFIFCFKLLYF 1423
Db 1364 IVDYEEPRROAVRSYNSRVKLSGDLCSVOGCRPCEDGASGSHHSVIFIFCFKLLYF 1423
Qy 1424 MELMWL 1428
Db 1424 MELMWL 1428
Qy 1441 MELMWL 1445
Db 1441 MELMWL 1445

```

```

ID ABB82168 standard; Protein; 1445 AA.
XX
AC ABB82168;
XX
DT 23-DEC-2002 (first entry)
XX
DE Human CD109 KI-H7 variant protein.
XX
KW CD109; alpha2 macroglobulin; thioester; cerebroprotective; cardiac;
KW immunosuppressive; haemostatic; anticoagulant; thrombolytic; human;
KW cardiovascular; vasotropic; gene therapy; CD109 KI-H7; variant.
XX
OS Homo sapiens.
XX
PN WO200270696-A2.
XX
PD 12-SEP-2002.
XX
PF 07-MAR-2002; 2002WO-CA00292.
XX
PR 07-MAR-2001; 2001US-273814P.
XX
PA (SCHU/) SCHUH A.
PA (SOTH/) SUTHERLAND R D.
PI Schuh A, Sutherland RD;
XX
DR WPI: 2002-713450/77.
DR N-PSDB; ABQ79967.
XX
PT New CD109 nucleic acids and polypeptides, useful in gene therapy,
PT particularly for treating strokes, myocardial infarctions, thrombosis,
PT thrombocytopenia, autoimmune diseases, or organ or bone marrow
PT transplantation -
XX
PS Claim 21; Fig 3b; 156pp; English.
XX
SS The invention relates to isolated nucleic acid molecules encoding CD109
SS polypeptides. These nucleic acid molecules include the human cDNA
SS sequences comprising CD109 KI, CD109 KI-H7, CD109 KI5 or their variants.
SS CD109 is a novel member of the alpha2 macroglobulin (alpha2M/C3, C4, C5
SS family of thioester-containing proteins. The CD109 polypeptides can be
SS expressed by standard recombinant methodology. The CD109 nucleic acid,
SS CD109 polypeptide or its mimetic, or the anti-CD109 antibody is useful
SS for treating or detecting a disease or disorder, e.g. conditions
SS associated with endothelial activation, platelet activation, activation
SS of the coagulation or fibrinolytic systems, activation of T lymphocytes
SS and of the complement system, quantitative or qualitative abnormalities
SS of platelet function, increased or impaired activation of the coagulation and/or
SS fibrinolytic systems, or impaired or increased immune activation. These
SS are also useful for treating cardiovascular disorders, stroke, myocardial
SS infarction, thrombosis, embolism, peripheral vascular disease,
SS thrombocytopenia, thrombocytopenia, autoimmune diseases, organ
SS transplantation, or bone marrow transplantation. The CD109 nucleic acids
SS or their antisense nucleotide sequence are useful in gene therapy for
SS treating these conditions. The present sequence represents the human
SS CD109 KI-H7 variant sequence.
XX
SQ Sequence 1445 AA;
XX
Query Match 99.6%; Score 7317.5; DB 23; Length 1445;
Best Local Similarity 98.7%; Pred. No. 0;
Matches 1426; Conservative 1; Mismatches 1; Indels 17; Gaps 1;
Qy 1 MCGPPLLTAAHLVCCTAALAVAPGPRFLVTAGIIRPGNVITIGVLLHCHBSQVTVKA 60
Db 1 MCGPPLLTAAHLVCCTAALAVAPGPRFLVTAGIIRPGNVITIGVLLHCHBSQVTVKA 60
Qy 61 ELKKTASNLTVSVLEAGVGEKSGFKTLTLPSPINSADITIELRYTGRQODELFSNST 120
Db 61 ELKKTASNLTVSVLEAGVGEKSGFKTLTLPSPINSADITIELRYTGRQODELFSNST 120

```

QY 121 RLSFTKRI SVFIQTDKALYKPKQVYKRIYVLTJESDFPKYKTSNLILKDKPKSNILQOML 180  
 Db 121 RLSFTKRI SVFIQTDKALYKPKQVYKRIYVLTJESDFPKYKTSNLILKDKPKSNILQOML 180  
 QY 181 SQQSLGVYISKTFOLSHPIIGDMSI QOVVNDQTYQSSQVSEYVLPRPEVLTQPLKCS 240  
 Db 181 SQQSLGVYISKTFOLSHPIIGDMSI QOVVNDQTYQSSQVSEYVLPRPEVLTQPLKCS 240  
 QY 241 MNSKILNGITITAKYTYGKPKVGDVLTJLPLSPFKSKKNIITKTFKINGSANFSPDEBMK 300  
 Db 241 MNSKILNGITITAKYTYGKPKVGDVLTJLPLSPFKSKKNIITKTFKINGSANFSPDEBMK 300  
 QY 301 NMVSSNGI SEYLDJSSPGPVEIITTTESVTGISRNSTNVFPKQDHYIIEFDYTVL 360  
 Db 301 NMVSSNGI SEYLDJSSPGPVEIITTTESVTGISRNSTNVFPKQDHYIIEFDYTVL 360  
 QY 361 KPSLFTATVKTADGNQLTLEBRNNVITVQNTTEVWGSNGSQKQKQKAVQKIN 420  
 Db 361 KPSLFTATVKTADGNQLTLEBRNNVITVQNTTEVWGSNGSQKQKQKAVQKIN 420  
 QY 421 TVPQSGTFKIEPPIEDSSELQKAYFLGSKSSMAVHSIFKSPKTYIOLKTRDENIKV 480  
 Db 421 TVPQSGTFKIEPPIEDSSELQKAYFLGSKSSMAVHSIFKSPKTYIOLKTRDENIKV 480  
 QY 481 SPFELVSGNKRKLKELSYNVVSRGQLVAVGKONSTMFSLTPBNSWTPKACVIVYIIEBDG 540  
 Db 481 SPFELVSGNKRKLKELSYNVVSRGQLVAVGKONSTMFSLTPBNSWTPKACVIVYIIEBDG 540  
 QY 541 EISIVLKI PVOLVKNKIKLWMSKYKAPSEKVSILRISVTPDPSI VGIIVANDKSVNLMN 600  
 Db 541 EISIVLKI PVOLVKNKIKLWMSKYKAPSEKVSILRISVTPDPSI VGIIVANDKSVNLMN 600  
 QY 601 ASNDITMENVVELELYNTGYLGMFMNSFAVFOCGMVLTDALTKDYIDGVYDNAY 660  
 Db 601 ASNDITMENVVELELYNTGYLGMFMNSFAVFOCGMVLTDALTKDYIDGVYDNAY 660  
 QY 661 AERFMEENEGHIVDIDHPSLGSSPHVRKHPETWIMLDTNMGYRYOEFEVTPDPSITSM 720  
 Db 661 AERFMEENEGHIVDIDHPSLGSSPHVRKHPETWIMLDTNMGYRYOEFEVTPDPSITSM 720  
 QY 721 VATGVISEDGLGLTTTPVELQAFOPFFIPLNLPYSVIRGSEFALETIENYLKDATEV 780  
 Db 721 VATGVISEDGLGLTTTPVELQAFOPFFIPLNLPYSVIRGSEFALETIENYLKDATEV 780  
 QY 781 KVIIEKSDKEDILMTSSEINATGHOOITLVSEDAATLFPRIPLHGBIPITWTALSP 840  
 Db 781 KVIIEKSDKEDILMTSSEINATGHOOITLVSEDAATLFPRIPLHGBIPITWTALSP 840  
 QY 841 ASDAVTOMILVYKABIEKSYOSIILDLTDNRLQSTLKTLSFSPNTVTGSEKRVQITAI 900  
 Db 841 ASDAVTOMILVYKABIEKSYOSIILDLTDNRLQSTLKTLSFSPNTVTGSEKRVQITAI 900  
 QY 901 GCVLGPISINGLASLIRMPYGCCEQNMNINFAPIYIILDYLTKKKQTLMDLKEKALSFMQ 960  
 Db 901 GCVLGPISINGLASLIRMPYGCCEQNMNINFAPIYIILDYLTKKKQTLMDLKEKALSFMQ 960  
 QY 961 YORELLYQREOSFSAFGNYDPSGSTWLSAFLRCLEADPYIDIDQNVLAHTYTWLKG 1020  
 Db 961 YORELLYQREOSFSAFGNYDPSGSTWLSAFLRCLEADPYIDIDQNVLAHTYTWLKG 1020  
 QY 1021 QKSNGEFPDGRVISELQGNKSPVTLTAIVTSLGGRKQPNIDVOESIHFLSEFS 1080  
 Db 1021 QKSNGEFPDGRVISELQGNKSPVTLTAIVTSLGGRKQPNIDVOESIHFLSEFS 1080  
 QY 1081 RGISNNTYLTALITVALSVGSPKAKALNMLTWRAEOBQGMQFWVSSEKLSDSWQPSL 1140  
 Db 1081 RGISNNTYLTALITVALSVGSPKAKALNMLTWRAEOBQGMQFWVSSEKLSDSWQPSL 1140  
 QY 1141 DIEVAAYALLSHPLQOTSEGIPIRMWLSRQNSIGCPASTODTTVALKALSEFALNMT 1200  
 Db 1141 DIEVAAYALLSHPLQOTSEGIPIRMWLSRQNSIGCPASTODTTVALKALSEFALNMT 1200  
 QY 1201 ERTNIQVTVTGPSSPSP-----LAVVQPMVANSANGFPAICQLNV 1243

Db 1201 ERTNIQVTVTGPSSPSPFKFLIDTHNRLLOTAELAVVQPMVANSANGFPAICQLNV 1260  
 QY 1244 YNVKASGSSRRRRSIOQOEAFDLVAKENDDINHDVLAVNCSFSGPGRSGMALMEVNL 1303  
 Db 1261 YNVKASGSSRRRRSIOQOEAFDLVAKENDDINHDVLAVNCSFSGPGRSGMALMEVNL 1320  
 QY 1304 LSGFMVSEALISLETYKVEYDHGKNTLYLDSVNETQFCVNIPIAVNFKVSNTOQASVS 1363  
 Db 1321 LSGFMVSEALISLETYKVEYDHGKNTLYLDSVNETQFCVNIPIAVNFKVSNTOQASVS 1380  
 QY 1364 IVDYIEPRRQAVRSYNSVVKLSGDLCSDVQGPCPCEDGASGSHHSVIFIFCPKLIYF 1423  
 Db 1381 IVDYIEPRRQAVRSYNSVVKLSGDLCSDVQGPCPCEDGASGSHHSVIFIFCPKLIYF 1440  
 QY 1424 MELWL 1428  
 Db 1441 MELWL 1445

RESULT 8  
 AA019373  
 ID AA019373 standard; Protein; 1445 AA.  
 XX AA019373;  
 AC AA019373;  
 XX  
 DT 04-DEC-2002 (first entry)  
 XX  
 DE Human platelet alloantigen Govb.  
 XX  
 KW Human; platelet alloantigen; Gova; Govb; single nucleotide polymorphism;  
 KW SNP; biallelic; bleeding disorder; post-transfusion purpura;  
 KW post-transfusion platelet refractoriness; haemostatic; vaccine;  
 KW neonatal alloimmune thrombocytopenia.  
 XX  
 OS Homo sapiens.  
 XX  
 PN MO20020738-A2.  
 XX  
 PD 12-SEP-2002.  
 XX  
 PF 07-MAR-2002; 2002WO-CA00291.  
 XX  
 PR 07-MAR-2001; 2001US-273941P.  
 XX  
 PA (SCHU/) SCHUH A.  
 XX  
 PI Schuh A, Ouwehand W;  
 XX  
 DR MPI; 2002-713460/77.  
 XX  
 DR N-PSDB; AAL49816.  
 XX  
 PT New isolated oligonucleotide binding to a region of CD109 nucleic acid  
 PT having a single nucleotide polymorphism that distinguishes a Gova  
 PT and/or Govb allele, useful for treating blood disorders e.g. alloimmune  
 PT thrombocytopenia  
 XX  
 PS Disclosure; Page 41-47; 69pp; English.  
 XX  
 CC The present invention relates to a sequence capable of binding  
 CC specifically to a CD109 nucleic acid which has a single nucleotide  
 CC polymorphism that distinguishes the Gova and Govb alleles. Detection of  
 CC the Gov genotype is useful for detecting whether the subject has or is at  
 CC risk of a blood disease, disorder or abnormal physical state, such as  
 CC bleeding, or increased risk of bleeding, due to alloimmune destruction of  
 CC blood platelets, e.g., post-transfusion purpura, post-transfusion  
 CC platelet refractoriness or neonatal alloimmune thrombocytopenia. The  
 CC nucleic acid and polypeptide are useful for Gov genotyping or phenotyping  
 CC individuals. The present sequence is the Govb encoded protein.  
 XX  
 SO Sequence 1445 AA;  
 XX

Query Match 99.6%; Score 7317.5; DB 23; Length 1445;

Best Local Similarity 98.7%; Pred. No. 0;  
Matches 1426; Conservative 1; Mismatches 17; Gaps 1;

Qy	1	MOGPBLTAAHLCTCTAALAVAPBRPLTAPGIIIRGGAVTLIGVLELHCPQVYTKA	60
Db	1	MOGPBLTAAHLCTCTAALAVAPBRPLTAPGIIIRGGAVTLIGVLELHCPQVYTKA	60
Qy	61	ELLKTASNLTVSVLEAEGVFEKGSFKTLTLPSLPLNSADEIYELAVTGRTODELFNSNT	120
Db	61	ELLKTASNLTVSVLEAEGVFEKGSFKTLTLPSLPLNSADEIYELAVTGRTODELFNSNT	120
Qy	121	RLSFETKRIISVFIQTDKALKYRKQEVKRIIYTLBSDFPKPYKTSNLILIKDKPSNLIQOML	180
Db	121	RLSFETKRIISVFIQTDKALKYRKQEVKRIIYTLBSDFPKPYKTSNLILIKDKPSNLIQOML	180
Qy	181	SOQSDLGVIKSTFOLSSHPIILGDMSIQOVANDQTYGQSFQVSEYVLPFEVTLQPLKCS	240
Db	181	SOQSDLGVIKSTFOLSSHPIILGDMSIQOVANDQTYGQSFQVSEYVLPFEVTLQPLKCS	240
Qy	241	MNSKRLNGTITAKYTYGKPVKGDVTLJTEPLSPFWGKKNNITKTPKINGSANFSFDEBEMK	300
Db	241	MNSKRLNGTITAKYTYGKPVKGDVTLJTEPLSPFWGKKNNITKTPKINGSANFSFDEBEMK	300
Qy	301	NVMOSSNGLSEYLDLSSPGPVEILITTTVESTGISRNVSTNVFFKQOHYIIEFPDYTVL	360
Db	301	NVMOSSNGLSEYLDLSSPGPVEILITTTVESTGISRNVSTNVFFKQOHYIIEFPDYTVL	360
Qy	361	KPSLNFATVATVATRADGNQTLTEBRNNVVITVQORANTEWMSGNSNGOMKMEAVQKNTY	420
Db	361	KPSLNFATVATVATRADGNQTLTEBRNNVVITVQORANTEWMSGNSNGOMKMEAVQKNTY	420
Qy	421	TVPOSGTEKIEFPPILEDSESELQKAYFLGSKSSMAVHSLFKSPSKTYIQLKTRDENIKVG	480
Db	421	TVPOSGTEKIEFPPILEDSESELQKAYFLGSKSSMAVHSLFKSPSKTYIQLKTRDENIKVG	480
Qy	481	SPFELVYSGNRKLKELSTYVVSRRGLVAVGKONSIMESLTJEBNSMTPRACVITYIIEEDG	540
Db	481	SPFELVYSGNRKLKELSTYVVSRRGLVAVGKONSIMESLTJEBNSMTPRACVITYIIEEDG	540
Qy	541	EIISDVLKIPOLVFNKIKLYMSFKVAKPESEKVSRLSTVQPSDIBGIVAVDVKSVNLNM	600
Db	541	EIISDVLKIPOLVFNKIKLYMSFKVAKPESEKVSRLSTVQPSDIBGIVAVDVKSVNLNM	600
Qy	601	ASNDITMENNVHELELYNTGYLLGFMFMSFAVQEGCGLMVLTDANLTQDYIDGVYDNAEY	660
Db	601	ASNDITMENNVHELELYNTGYLLGFMFMSFAVQEGCGLMVLTDANLTQDYIDGVYDNAEY	660
Qy	661	AERFPEENEGHIVDHDFFSLGSSSPVYRGFPETWMLDTNMGCRYQGFVTVVDSITSM	720
Db	661	AERFPEENEGHIVDHDFFSLGSSSPVYRGFPETWMLDTNMGCRYQGFVTVVDSITSM	720
Qy	721	VATGFVISEDGLGLTTTPVELQAFQPFIFLNLPSYVIRGEEFALETIENYLKDATEV	780
Db	721	VATGFVISEDGLGLTTTPVELQAFQPFIFLNLPSYVIRGEEFALETIENYLKDATEV	780
Qy	781	KVIEEKSDKFDILMTSSSINATGHQOTLVPSBDGATVLPPIRPHLGEIPITYTASLPT	840
Db	781	KVIEEKSDKFDILMTSSSINATGHQOTLVPSBDGATVLPPIRPHLGEIPITYTASLPT	840
Qy	841	ASDAVTQMLIKABEIEKSYQSILDLTDNRLOSTLTKTSSFPSPPNVTGSSERVQITAI	900
Db	841	ASDAVTQMLIKABEIEKSYQSILDLTDNRLOSTLTKTSSFPSPPNVTGSSERVQITAI	900
Qy	901	GDVAGPSINGLASLIRMPYGCGEQNMIFAPNIYLDLYTKKKQLTJDKKALSPFMQ	960
Db	901	GDVAGPSINGLASLIRMPYGCGEQNMIFAPNIYLDLYTKKKQLTJDKKALSPFMQ	960
Qy	961	YORELLYOREDGSFSAFNGNDPSSGTLVSAFVLRGCEADVPYIDIDQNVLRHTYTWLKGH	1020
Db	961	YORELLYOREDGSFSAFNGNDPSSGTLVSAFVLRGCEADVPYIDIDQNVLRHTYTWLKGH	1020
Qy	1021	OKSNGEFMDPRVLIHSELQGNKSPVTLTAYTVSLGLGRKYOPRINDVQESITHLESFS	1080
Db	1021	OKSNGEFMDPRVLIHSELQGNKSPVTLTAYTVSLGLGRKYOPRINDVQESITHLESFS	1080

Db	1021	OKSNSEFMDPCRVHSELQGNKSPVTLTAVIVTSLGGRKYQPNIDVQSIHFLESEFS	1089
Qy	1081	RGISDNTTALLITYALSVSQSPRAKELNMLTWRAEGEGMGMWVSESLSPSWQPSRL	1140
Db	1081	RGISDNTTALLITYALSVSQSPRAKELNMLTWRAEGEGMGMWVSESLSPSWQPSRL	1140
Qy	1141	DIEVAAYALLSHFLQFQTSSEGIPIRMWLSHRQNSLGGFASSTQDTTVALKALSEFALMNT	1200
Db	1141	DIEVAAYALLSHFLQFQTSSEGIPIRMWLSHRQNSLGGFASSTQDTTVALKALSEFALMNT	1200
Qy	1201	ERTNIQVTVTGSSPSP-----LAVQPMVAVNISANGFPAICQLNV	1243
Db	1201	ERTNIQVTVTGSSPSP-----LAVQPMVAVNISANGFPAICQLNV	1260
Qy	1244	YNVAKSSSSRRRTIQOQEAFDLDVAVKMKDLDNHDVLCVCSFGSPGSGMALMEVNL	1303
Db	1261	YNVAKSSSSRRRTIQOQEAFDLDVAVKMKDLDNHDVLCVCSFGSPGSGMALMEVNL	1320
Qy	1304	LSGFMVSEBALSLSETYKAKVEYDHGKLNLYLDSVNETQFCVNI PAVANFKAUSTQDASVS	1363
Db	1321	LSGFMVSEBALSLSETYKAKVEYDHGKLNLYLDSVNETQFCVNI PAVANFKAUSTQDASVS	1380
Qy	1364	IVDYEBPRQAVRSYNSVXKLSGCDLCSDVQGRPCEDGASGSHHSSVIFIFCFKLLYF	1423
Db	1381	IVDYEBPRQAVRSYNSVXKLSGCDLCSDVQGRPCEDGASGSHHSSVIFIFCFKLLYF	1440
Qy	1424	MELMW 1428	
Db	1441	MELMW 1445	
RESULT 9			
AA12127			
ID	AA12127	standard; Protein; 1445 AA.	
XX	AA12127;		
XX	02-FEB-2001	(first entry)	
DE	Hydrophobic domain protein isolated from HT-1080 cells.		
XX	Human; secreted protein; membrane protein; hydrophobic domain;		
KW	proliferation control; differentiation induction; material transport;		
KM	biophysics; signal receptor; ion channel; transporter; immunostimulant;		
KW	immunobupressant; haematopoiesis regulator; chemotactic; chemokinetic;		
KM	haemostatic; thrombolytic; anti-inflammatory; tumour inhibition;		
XX	autoimmune disease; Alzheimer's disease; Parkinson's disease; cancer.		
OS	Homo sapiens.		
XX	WO200029448-A2.		
XX	25-MAY-2000.		
XX	17-NOV-1999;	99MO-JP06412.	
XX	17-NOV-1998;	98JP-0326255.	
PR	22-DEC-1998;	98JP-0364315.	
PR	16-MAR-1999;	99JP-0069811.	
PR	27-APR-1999;	99JP-0119299.	
PR	19-MAY-1999;	99JP-0138169.	
XX	(SAGA ) SAGAMI CHEM RES CENT.		
PA	(PROT-) PROTEGENE INC.		
XX	Kato S, Kimura T;		
XX	WI: 2000-387753/33.		
DR	N-PSDB; AAA60199, AAA62010.		
PT	Proteins comprising hydrophobic regions, such as secretory and membrane		
PT	proteins, useful in research and diagnostics and having various		
PT	activities e.g. immunomodulatory, antiinflammatory, chemokinetic,		

PT proteins comprising hydrophobic regions, such as secretory and membrane proteins, useful in research and diagnostics and having various activities e.g. immunomodulatory, antiinflammatory, chemokinetic, PT



PN WO200285942-A2.  
 XX 31-OCT-2002.  
 XX 24-APR-2002; 2002WO-CA00560.  
 XX 24-APR-2001; 2001US-285713P.  
 PR 14-FEB-2002; 2002US-356163P.  
 XX (UMC-) UNIV MCGILL.  
 PA Philip A, Tam B;  
 PI WPI; 2003-093100/08.  
 DR N-PSDB; AAD49436.  
 XX  
 PT Novel transforming growth factor (TGF)-beta 1 binding reagent which  
 PT comprises r150 protein which acts as accessory receptor of TGF-beta,  
 PT useful for negatively modulating TGF-beta activity, and thus for  
 PT treating cancer.  
 PS Claim 1; Page 112-117; 127pp; English.  
 CC The invention relates to novel transforming growth factor (TGF)-beta1  
 CC binding reagent which comprises a glycosylphosphatidy1 inositol (GPI)-  
 CC anchored TGF-beta1 binding protein referred to as r150 which acts as  
 CC an accessory receptor of TGF-beta. The invention is used for negatively  
 CC modulating TGF-beta activity, and thus for treating conditions  
 CC characterised by overproduction of TGF-beta, such as cancer. Antisense  
 CC molecules of the invention are used for increasing TGF-beta availability  
 CC and increase graft success. The present sequence is human blood cell  
 CC surface antigen, CD109. CD109 is a variant of r150 sequence.  
 XX  
 SQ Sequence 1382 AA;

Query Match 96.4%; Score 7081; DB 24; Length 1382;  
 Best Local Similarity 99.8%; Pred. No. 0;  
 Matches 1379; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 22 VAPGRFLVTAAGIIRPGNVITIGVLEHPCQVTVKAEELKTASNLTVSLAEAGVFE 81  
 DB 1 VAPGRFLVTAAGIIRPGNVITIGVLEHPCQVTVKAEELKTASNLTVSLAEAGVFE 60  
 QY 82 KGSFKTLTLPSPINSADEIYELVGTODEILFNSSTRLSFEFKRISVFIQTDKALYK 141  
 DB 61 KGSFKTLTLPSPINSADEIYELVGTODEILFNSSTRLSFEFKRISVFIQTDKALYK 120  
 QY 142 PKQEVKFRIVTLFSDPKRYKTSNLIKIDPKSNLIQOMLSQOSDLGVISKTFOUSSHPIL 201  
 DB 121 PKQEVKFRIVTLFSDPKRYKTSNLIKIDPKSNLIQOMLSQOSDLGVISKTFOUSSHPIL 180  
 QY 202 GDMGSIQOVNDQTYQSFQVSEYVLPKFEVTLQTPVCSNMSKILNNTTAKTYYGKPVK 261  
 DB 181 GDMGSIQOVNDQTYQSFQVSEYVLPKFEVTLQTPVCSNMSKILNNTTAKTYYGKPVK 240  
 QY 262 GGVTLTFLPLSFQKKNITKTFKINGSANFSFNDDEKKNVMSNGLSEYLDLSGKPV 321  
 DB 241 GGVTLTFLPLSFQKKNITKTFKINGSANFSFNDDEKKNVMSNGLSEYLDLSGKPV 300  
 QY 322 EILTTVTESVTGISRNVTNVFKQHDYIIEFPDYTVLKPSLNFATVKTATRADGNQLT 381  
 DB 301 EILTTVTESVTGISRNVTNVFKQHDYIIEFPDYTVLKPSLNFATVKTATRADGNQLT 360  
 QY 382 LEEERNNVVITVQONTTEYSGNSGNQKMEAVOKINITYVPGSTFKIEPILEDSESL 441  
 DB 361 LEEERNNVVITVQONTTEYSGNSGNQKMEAVOKINITYVPGSTFKIEPILEDSESL 420  
 QY 442 QLKAYFLGSKSSMAVHSLFKSPSKTYIQLKTRDENIKVGSFPELVVSGNKKLKELSYVV 501  
 DB 421 QLKAYFLGSKSSMAVHSLFKSPSKTYIQLKTRDENIKVGSFPELVVSGNKKLKELSYVV 480  
 QY 502 SRGQLVAVGKONSTMFSLTPENSWTPKACVIVYIIEDDGEIISDVLKIPQVLFKNKIKL 561  
 DB 561 SRGQLVAVGKONSTMFSLTPENSWTPKACVIVYIIEDDGEIISDVLKIPQVLFKNKIKL 540

DB 481 SRGQLVAVGKONSTMFSLTPENSWTPKACVIVYIIEDDGEIISDVLKIPQVLFKNKIKL 540  
 QY 562 YMSKVAAEPESEKSLRISVTPQDSIVGIVAVDKSNVLMAASNDITMENVHELELYNTGY 621  
 DB 541 YMSKVAAEPESEKSLRISVTPQDSIVGIVAVDKSNVLMAASNDITMENVHELELYNTGY 600  
 QY 622 YLGMFNNSEFAVFOECGLMWLTDLNLTQDYIDGYDNAEVAERMEENEGHIVDIHDSFLG 681  
 DB 601 YLGMFNNSEFAVFOECGLMWLTDLNLTQDYIDGYDNAEVAERMEENEGHIVDIHDSFLG 660  
 QY 682 SSFHVAKHPEETIMWLTQNGVRIYQEFETVPDSITSWATGTFVISEDGLGTLTPVE 741  
 DB 661 SSFHVAKHPEETIMWLTQNGVRIYQEFETVPDSITSWATGTFVISEDGLGTLTPVE 720  
 QY 742 LQAFQPFPIFLNPYSVIRGEFEALITTFNVLKQATEVAVIIEKSDKFDILMTSEINA 801  
 DB 721 LQAFQPFPIFLNPYSVIRGEFEALITTFNVLKQATEVAVIIEKSDKFDILMTSEINA 780  
 QY 802 TGHQOQTLVPSDEGATVLPFIRPTHGELPIYVYALSPTASDVAVTOMILYKAGIEKSY 861  
 DB 781 TGHQOQTLVPSDEGATVLPFIRPTHGELPIYVYALSPTASDVAVTOMILYKAGIEKSY 840  
 QY 862 QSTILLDTNRLQSTLKTLSFSFPNTVMSSEBVOITAGDVLGSPINGLASLIRMPYGC 921  
 DB 841 QSTILLDTNRLQSTLKTLSFSFPNTVMSSEBVOITAGDVLGSPINGLASLIRMPYGC 900  
 QY 922 GEONMINFAPNIYILDYLTQKQTLNLEKALSPRQGYORELLYQREDSFSARGNYD 981  
 DB 901 GEONMINFAPNIYILDYLTQKQTLNLEKALSPRQGYORELLYQREDSFSARGNYD 960  
 QY 982 PGSGTWLSAFVLCFLEADPYIDIDQNVLHRTYTWLKHQKNGEFPDGRVHSELQGG 1041  
 DB 961 PGSGTWLSAFVLCFLEADPYIDIDQNVLHRTYTWLKHQKNGEFPDGRVHSELQGG 1020  
 QY 1042 NKSPMTLTATYVSLGCRYQPINVOESIHRLSEPSRGISDNTLTALITALSYSVS 1101  
 DB 1021 NKSPMTLTATYVSLGCRYQPINVOESIHRLSEPSRGISDNTLTALITALSYSVS 1080  
 QY 1102 PKAKEALNMLTWRAEOEGQMFVSSSEKLSDSQORSIDIEVAAYALSHFLQFQTSSEG 1161  
 DB 1081 PKAKEALNMLTWRAEOEGQMFVSSSEKLSDSQORSIDIEVAAYALSHFLQFQTSSEG 1140  
 QY 1162 IPIRMWLSRQNSLGGPASTQDTTVALKALSEPAALMTERINIQTVTGSSSPSLAVV 1221  
 DB 1141 IPIRMWLSRQNSLGGPASTQDTTVALKALSEPAALMTERINIQTVTGSSSPSLAVV 1200  
 QY 1222 QPMANVISANGEPFALCOLNVVYNVYKASGSSRRRSIQNOEAPDLVAVKENDDLNHYD 1281  
 DB 1201 QPMANVISANGEPFALCOLNVVYNVYKASGSSRRRSIQNOEAPDLVAVKENDDLNHYD 1260  
 QY 1282 LNVCTSPSGRSGMALMEVNLDSGFMPVSEALISLSEYVKKVEYDHGKTLNLYDSVNETQ 1341  
 DB 1261 LNVCTSPSGRSGMALMEVNLDSGFMPVSEALISLSEYVKKVEYDHGKTLNLYDSVNETQ 1320  
 QY 1342 FCVNIIPAVNFKVSNTOASVSIVDYIEBPRQAVRSYNSVEVKLSGCDLCSGVGCRPCE 1401  
 DB 1321 FCVNIIPAVNFKVSNTOASVSIVDYIEBPRQAVRSYNSVEVKLSGCDLCSGVGCRPCE 1380  
 QY 1402 GA 1403  
 DB 1381 GA 1382

RESULT 11  
 AAE32015  
 ID AAE32015 standard; Protein; 1399 AA.  
 XX AAE32015;  
 AC AAE32015;  
 XX 24-MAR-2003 (first entry)  
 XX Human r150 protein #3.  
 XX





RESULT 12  
AB82169  
ID ABB82169 standard; Protein; 1067 AA.  
XX  
AC ABB82169;  
XX  
DT 23-DEC-2002 (first entry)  
DE Human CD109 K15 protein.  
XX  
KW CD109: alpha2 macroglobulin; thioester; cerebroprotective; cardiac; immunosuppressive; haemostatic; anticoagulant; thrombolytic; human; cardiovascular; vasotropic; gene therapy; CD109 K15.  
XX  
OS Homo sapiens.  
XX  
PN MO200270696-A2.  
XX  
PD 12-SEP-2002.  
XX  
PF 07-MAR-2002; 2002MO-CA00292.  
XX  
PR 07-MAR-2001; 2001US-273814P.  
XX  
PA (SCHU/) SCHUH A.  
XX  
PA (SUTH/) SUTHERLAND R D.  
XX  
PI Schuh A, Sutherland RD,  
XX  
DR WPI: 2002-713450/77.  
XX  
DR N-PSDB; ABQ79968.  
XX  
PT New CD109 nucleic acids and polypeptides, useful in gene therapy,  
PT particularly for treating strokes, myocardial infarctions, thrombosis,  
PT thrombocytopenia, autoimmune diseases, or organ or bone marrow  
PT transplantation -  
XX  
PS Claim 21; Fig 5a; 156pp; English.  
XX  
CC The invention relates to isolated nucleic acid molecules encoding CD109  
CC polypeptides. These nucleic acid molecules include the human cDNA  
CC sequences comprising CD109 K1, CD109 K1-H7, CD109 K15 or their variants.  
CC CD109 is a novel member of the alpha2 macroglobulin (alpha2M)/C3, C4, C5  
CC family of thioester-containing proteins. The CD109 polypeptides can be  
CC expressed by standard recombinant methodology. The CD109 nucleic acid,  
CC CD109 polypeptide or its mimetic, or the anti-CD109 antibody is useful  
CC for treating or detecting a disease or disorder, e.g. conditions  
CC associated with endothelial activation, platelet activation, activation  
CC of the coagulation or fibrinolytic systems, activation of T lymphocytes  
CC and of the complement system, quantitative or qualitative abnormalities  
CC of platelet function, increased or impaired platelet aggregation and  
CC activation, increased or impaired activation of the coagulation and/or  
CC fibrinolytic systems, or impaired or increased immune activation. These  
CC are also useful for treating cardiovascular disorders, stroke, myocardial  
CC infarction, thrombosis, embolism, peripheral vascular disease,  
CC thrombocytopenia, thrombocytopenia, autoimmune diseases, organ  
CC transplantation, or bone marrow transplantation. The CD109 nucleic acids  
CC or their antisense nucleotide sequence are useful in gene therapy for  
CC treating these conditions. The present sequence represents the human  
CC CD109 K15 sequence.  
XX  
SQ Sequence 1067 AA;

Query March 74.4%; Score 5470; DB 23; Length 1067;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 1062; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGGPPLTAALHLCVCTAALAVAPGPRFLVTAAGIIRPGNVITGVLELHCPQVTVKA 60  
DB 1 MGGPPLTAALHLCVCTAALAVAPGPRFLVTAAGIIRPGNVITGVLELHCPQVTVKA 60  
QY 61 ELKTAASNLTVSVLEAGVFEKGSFKTLTLPSPPLNSADEIYELRVGTODEILFNSNT 120

DB 61 ELKTAASNLTVSVLEAGVFEKGSFKTLTLPSPPLNSADEIYELRVGTODEILFNSNT 120  
QY 121 RLSFETKRISVFIQTDKALYKPKQEVKFRIVTLFSDPKPYKSLNLIIDPKSNLIQQWL 180  
DB 121 RLSFETKRISVFIQTDKALYKPKQEVKFRIVTLFSDPKPYKSLNLIIDPKSNLIQQWL 180  
QY 181 SQQSDLGIVISKTFQLSHPHILGWSIQOVNDQTYIQSFQVSEYVLPKFEVTLQTPLYCS 240  
DB 181 SQQSDLGIVISKTFQLSHPHILGWSIQOVNDQTYIQSFQVSEYVLPKFEVTLQTPLYCS 240  
QY 241 MNSKHNGTITTAQYTGKPRKGDVTLTFLPLSPWGGKXNITKPKFNGSANSFNDSEMK 300  
DB 241 MNSKHNGTITTAQYTGKPRKGDVTLTFLPLSPWGGKXNITKPKFNGSANSFNDSEMK 300  
QY 301 NWMDSSNGLSSEYDLSSPGVEILTTVTESVTGISHNVTSTNVEFKOHDIIEFFDYTVL 360  
DB 301 NWMDSSNGLSSEYDLSSPGVEILTTVTESVTGISHNVTSTNVEFKOHDIIEFFDYTVL 360  
QY 361 KPSLNFATATVKTTRADGNQTLTEERRNNVITYTONNYTEYWSGNSGNQKMAVOKINY 420  
DB 361 KPSLNFATATVKTTRADGNQTLTEERRNNVITYTONNYTEYWSGNSGNQKMAVOKINY 420  
QY 421 TVPQSGFKIEPFIILDSSELOKAYFLGSKSSMAVHSLFKSPSKTYIQLKTRDENIKVG 480  
DB 421 TVPQSGFKIEPFIILDSSELOKAYFLGSKSSMAVHSLFKSPSKTYIQLKTRDENIKVG 480  
QY 481 SPFELVVSNGKRLKELSYVWSRGQLAVAGKONSTWFSLTPEKSWTPKACVITYYIEDDG 540  
DB 481 SPFELVVSNGKRLKELSYVWSRGQLAVAGKONSTWFSLTPEKSWTPKACVITYYIEDDG 540  
QY 541 EITSDVLIKIPVQLVFNKIKLKYMSKYKABSEKVSIRISVTPDSIVGIVAVDKSVNLNN 600  
DB 541 EITSDVLIKIPVQLVFNKIKLKYMSKYKABSEKVSIRISVTPDSIVGIVAVDKSVNLNN 600  
QY 541 EITSDVLIKIPVQLVFNKIKLKYMSKYKABSEKVSIRISVTPDSIVGIVAVDKSVNLNN 600  
QY 601 ASNDITMENVHLELYNNGYVYLGEMFNSPFAVQECGLWLDANLTCKYIDGVYNAEY 660  
DB 601 ASNDITMENVHLELYNNGYVYLGEMFNSPFAVQECGLWLDANLTCKYIDGVYNAEY 660  
QY 661 AERFMESENGHYDIDHPSIGSSPHVRKHPETWIMLDTNMGRIYQEFVTVPDSITSW 720  
DB 661 AERFMESENGHYDIDHPSIGSSPHVRKHPETWIMLDTNMGRIYQEFVTVPDSITSW 720  
QY 721 VATGFIYSEDLGGLTTPVELOAFQPFIFLNLPSVIRGESEFALETITFNLYKDATEV 780  
DB 721 VATGFIYSEDLGGLTTPVELOAFQPFIFLNLPSVIRGESEFALETITFNLYKDATEV 780  
QY 781 KVIIKSDKRDILMTSEINATGHQOTLLVPSBDGATVLPPIRPTHGELPIVTVALSPT 840  
DB 781 KVIIKSDKRDILMTSEINATGHQOTLLVPSBDGATVLPPIRPTHGELPIVTVALSPT 840  
QY 841 ASDAVTOMILVAKEGIEKYSQSILDLTDNRLOSTLTKLSFSPPNVTGSEKVOITAI 900  
DB 841 ASDAVTOMILVAKEGIEKYSQSILDLTDNRLOSTLTKLSFSPPNVTGSEKVOITAI 900  
QY 901 GDVLGDSINGLASLIMPYGCGEONNINFAPNIYILDYITKKKQLTIDNKEKALSPMRQ 960  
DB 901 GDVLGDSINGLASLIMPYGCGEONNINFAPNIYILDYITKKKQLTIDNKEKALSPMRQ 960  
QY 961 YQRELLYQREDSFSAFGNYDPGSGTWLSAFVLRCLLEADPYIDIONVLAHRYTTLKGH 1020  
DB 961 YQRELLYQREDSFSAFGNYDPGSGTWLSAFVLRCLLEADPYIDIONVLAHRYTTLKGH 1020  
QY 1021 QKNGSEFMDPGRVIHSELQGNKSPVTLTAIYITSLGGRKYQ 1063  
DB 1021 QKNGSEFMDPGRVIHSELQGNKSPVTLTAIYITSLGGRKYQ 1063

RESULT 13  
AB82170  
ID ABB82170 standard; Protein; 1067 AA.  
XX  
AC ABB82170;  
XX



KW therapy; blood cell surface antigen, CD109.  
 XX Homo sapiens.  
 XX MO200285942-A2.  
 XX 31-OCT-2002.  
 XX 24-APR-2002; 2002MO-CA00560.  
 XX 24-APR-2001; 2001US-285713P.  
 PR 14-FEB-2002; 2002US-356163P.  
 XX (UYMC-) UNIV MCGILL.  
 PA Philip A, Tam B;  
 PI WPI; 2003-093100/08.  
 DR  
 XX Novel transforming growth factor (TGF)-beta 1 binding reagent which  
 PT comprises r150 protein which acts as accessory receptor of TGF-beta,  
 PT useful for negatively modulating TGF-beta activity, and thus for  
 PT treating cancer -  
 XX Disclosure; Fig 17; 127pp; English.  
 PS  
 XX The invention relates to novel transforming growth factor (TGF)-beta1  
 CC binding reagent which comprises a glycosylphosphatidylinositol (GPI)-  
 CC anchored TGF-beta binding protein referred to as r150 which acts as  
 CC an accessory receptor of TGF-beta. The invention is used for negatively  
 CC modulating TGF-beta activity, and thus for treating cancer. Antisense  
 CC characterised by overproduction of TGF-beta, such as cancer. Antisense  
 CC molecules of the invention are used for increasing TGF-beta availability  
 CC and increase graft success. The present sequence is human blood cell  
 CC surface antigen, CD109. CD109 is a variant of r150 sequence  
 CC Note: This sequence is stated to be encoded by human CD109 DNA shown  
 CC in Figure 17 of the specification, but this does not appear to be the  
 CC case.  
 XX  
 XX  
 SQ Sequence 1445 AA;  
 Query Match 38.9%; Score 2860.5; DB 24; Length 1445;  
 Best Local Similarity 59.9%; Pred. No. 3.3e-178;  
 Matches 861; Conservative 1; Mismatches 559; Indels 17; Gaps 1;

QY 421 TVPQSGTFKIEPPILEDSSSELDKAVFLGSKSSMAVHSLFKSPSKYTIQKTDENIKVG 480  
 DB 421 TVPQSGTFPLIGPPILGSSGLGLATPLGLSSMAVHSLPLSPSLTTIGLTAAGAILVG 480  
 QY 481 SPFEIVVSGNKRKLKEISYMWVSRGOLAVAGKQNSTPESLTPENSMWPKACVITYIEDDG 540  
 DB 481 SPPELVVSGALALLGLSTWVSRGOLAVAGLGNSTPESLTPENSMWPKACVITYIEDDG 540  
 QY 541 EISDVLKIPVQVFNKIKLWMSKYAESEKYSRLISVTPQDSIVGIYAVDKSVNLNN 600  
 DB 541 GIIISAVLLIPVGLPPLAILLITTSIVLAGSGVLSLAVSTGPASIVGIYAVDLVALMA 600  
 QY 601 ASNDITMENVHELELYNTGYIYGMFNSPAVPQEGGLVYLDANLTQYIDGYDNAEY 660  
 DB 601 ASNAITWGAHVHGLATGTGTMEMASPAVPGGGLTVLTAALLTATAGVTAAGT 660  
 QY 661 AERFMEBEGHIVDHFSLGSSPHVRKHPETWIMDTMWGRYIQEPFVTPDSTSW 720  
 DB 661 AGAPMGAGSHIVAHAPSLGSSPHVALHPGTTITLATMGPAITGGPGCVTPASTIST 720  
 QY 721 VATGFVISEDIGLGLTTTPVELQAFQPFIFLNLPSYVIRGEFPALEITTFNYLKDATEV 780  
 DB 721 VATGPVISGALGGLTTPVGLGAPGPPIPLALPTSVIAGGPGALGITTPATLLAATGV 780  
 QY 781 KVIEKSDKFDIIMTSSEINATCHQOTLVPSBDGATVLPPIPTHGEPITVTALSPT 840  
 DB 781 LVIIIGLSALPAILMTSAGIATAGHGGTLVPSGAGATVLPPIPTHGEGIPITVTALSPT 840  
 QY 841 ASNAVQMLIVKAEIGKYSOSIILDTNDRQSLTKLSFEPNTVGSRRVQITAI 900  
 DB 841 ASNAVGMILVLAGGIGLSTSGSILALTAALAGSTLLTSPSPPAVTVGSAVGTAI 900  
 QY 901 GDTVLSGINSGLALIMFAPYCCGEONMIFAPNIYIIDLTKKQLOTDNLKEKALFEWROG 960  
 DB 901 GATVLSGINSGLALIMFAPYCCGEONMIFAPNIYIIDLTKKQLOTDNLKEKALFEWROG 960  
 QY 961 YORELLYOREDSFSFAGNYDPGSGTWSAFVLRCPLEADPYIDQNVLHRTYTWLKGH 1020  
 DB 961 TGAGLLTGAGAGSPASGATAPSGSTLLSAPVLACPLGAPTAIAGAVLHAATTTLLGH 1020  
 QY 1021 QKNGEFWDPGRYIHSELOGNKSPTLTATYITSLIGYKQVBNIDVQSHIFLSEPS 1080  
 DB 1021 GLSAGGPTAGAVIHSGLGAGLSPVLTATVITSLGTLATGPALAVGSHIPLSGPS 1080  
 QY 1081 RGISDNYTLLATVYALSSVSGPKAKALNMLTVRAEOEGMQPWSSEKSLDSWQPSRL 1140  
 DB 1081 AGISAAITLALITLALSSVSGPLALGALMLTTAAGCGGCMGPTVSSGSLSSSTGPASL 1140  
 QY 1141 DIEVAAYALLSHFLQOTSEGIPIKMWLSRQNSLGGPASTQDTYVALKALSEFALMNT 1200  
 DB 1141 AIGVAAATALLSHPLGCTSGGIPIMATLSGAASLGGPASTGATVYALLALSGPALMAT 1200  
 QY 1201 ERTNIQVTVGSSPSR-----LAVQPMANVISANGRFACQLNV 1243  
 DB 1201 GATTAIGTVTVGSSPSRPLPLATHAALLGTAGLAVGPMVAALISAAGGPATIGLAV 1260  
 QY 1244 YNFKAGSSRRRSIQNOEAFDDVAVKENKODLNVDTNCTSFSGPGRSGALMEVNL 1303  
 DB 1261 TATLASGSSAAASIGAGAPALAVLALALALAVLAVCTSPGPGASGALMGVAL 1320  
 QY 1304 LSGFMVPSAISLSETVKRYEDHGLNLVYDSVNETQPCVNIPAVNFKVSTQDASVS 1363  
 DB 1321 LSGPMVPSAISLSTGVLLVGTAGHLLATLLASVAGTCPCVAPPAVAPLVSATGAASVS 1380  
 QY 1364 IVUYIEPRQAVSYNSEVYKLSGCDLCSVQGRPREDDASGSHHSVFIICFLL 1421  
 DB 1381 IVATTGPAGAVASTASGVLLSSCALCSAVGAPCGAGASGSHHSVFIICPPLL 1438  
 RESULT 15  
 AAB50676  
 ID AAB50676 standard; Protein; 1508 AA.  
 XX



**Tue Jan 20 08:47:22 2004**

us-10-020-095-4.rag

Page 18

D<sub>b</sub> 1356 CCNTRPCKSMALAEIDALSGRPDAGVHTLSIEDLQREVMEXKDDTKKNYFNPZLG 1415  
Q<sub>y</sub> 1340 TQFCVNIPAVNFKKYSNTQDASVSIVDYEPBRQAVRFSYNSEVLSSCDLCSVDQGGRP 1398  
Db 1416 RPVCGLSTSLSDVTYGADQDKPANFRIVDYDDDEEQLKMITYAAKQRISJQEKG--EDCMP 1472

Search completed: January 15, 2004, 18:12:20  
Job time : 59 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 15, 2004, 18:11:15 ; Search time 22 Seconds

(without alignments)  
2746.359 Million cell updates/sec

Title: US-10-020-095-4

Perfect score: 7348

Sequence: 1 MGGPPLTAHLVCVCTAL.....HSSVIFCFKLLYEMELML 1428

Scoring table: BLOSUM62

Searched: 328717 segs, 42310958 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1441	19.6	1474	US-09-241-606-2	Sequence 2, Appl1
2	973.5	13.2	1651	US-08-447-411-2	Sequence 2, Appl1
3	931	12.7	1663	US-08-793-126-1	Sequence 1, Appl1
4	931	12.7	1663	US-09-132-271-1	Sequence 1, Appl1
5	931	12.7	1663	US-09-142-334-22	Sequence 22, Appl1
6	921	12.5	1676	US-08-487-283A-2	Sequence 2, Appl1
7	894	12.2	1642	US-08-662-227-2	Sequence 2, Appl1
8	894	12.2	1642	US-09-017-947-2	Sequence 2, Appl1
9	885	12.0	1642	US-08-447-411-45	Sequence 2, Appl1
10	880	12.0	1648	US-08-662-227-35	Sequence 35, Appl1
11	880	12.0	1648	US-09-017-947-35	Sequence 35, Appl1
12	768.5	10.5	1333	US-08-447-411-76	Sequence 76, Appl1
13	758.5	10.3	1333	US-08-662-227-34	Sequence 34, Appl1
14	758.5	10.3	1333	US-07-779-1947-34	Sequence 34, Appl1
15	739.5	10.3	1333	US-09-017-947-34	Sequence 34, Appl1
16	739.5	10.3	1333	US-09-132-271-1	Sequence 1, Appl1
17	215.5	2.9	250	US-09-241-606-4	Sequence 4, Appl1
18	185	2.5	183	US-09-311-352B-4	Sequence 4, Appl1
19	180	2.4	1341	US-09-252-991A-26785	Sequence 26785, A
20	174.5	2.4	2366	US-08-480-604A-10	Sequence 10, Appl1
21	174.5	2.4	2366	US-08-405-496A-10	Sequence 10, Appl1
22	174.5	2.4	2366	US-08-915-136-10	Sequence 10, Appl1
23	174.5	2.4	2366	US-08-957-310-10	Sequence 10, Appl1
24	174.5	2.4	2366	US-10-011-346-10	Sequence 10, Appl1
25	171	2.3	3290	US-09-328-352-5486	Sequence 5486, Ap
26	161	2.2	2199	US-08-793-273C-2	Sequence 2, Appl1
27	161	2.2	2199	PCT-US95-11684-2	Sequence 2, Appl1

28	160.5	2.2	403	US-08-118-674-1	Sequence 1, Appl1
29	156	2.1	10182	US-09-134-001C-3159	Sequence 3159, Ap
30	153.5	2.1	3169	US-09-453-702B-257	Sequence 257, App
31	152.5	2.1	1095	US-09-206-942-69	Sequence 69, Appl
32	152.5	2.1	1536	US-08-038-682-2	Sequence 2, Appl1
33	152.5	2.1	1536	US-08-302-832-2	Sequence 2, Appl1
34	152.5	2.1	1536	US-08-530-198-2	Sequence 2, Appl1
35	152.5	2.1	1536	US-08-469-880-2	Sequence 2, Appl1
36	152.5	2.1	1536	US-08-728-470-2	Sequence 2, Appl1
37	152.5	2.1	1536	US-08-617-697-2	Sequence 2, Appl1
38	152.5	2.1	1536	US-08-719-641-2	Sequence 2, Appl1
39	152.5	2.1	1536	US-09-206-942-67	Sequence 67, Appl
40	152.5	2.1	1844	US-08-851-567B-53	Sequence 53, Appl
41	152.5	2.1	2504	US-08-851-567B-12	Sequence 12, Appl
42	150.5	2.0	2516	US-08-851-567B-47	Sequence 47, Appl
43	148	2.0	2071	US-09-415-522-6	Sequence 6, Appl1
44	148	2.0	4563	US-09-108-006C-1	Sequence 1, Appl1
45	146.5	2.0	1391	US-09-106-568E-8	Sequence 8, Appl1

## ALIGNMENTS

RESULT 1	US-09-241-606-2	Sequence 2, Appl1	US/09241606
Patent No. 6472140			
GENERAL INFORMATION:			
APPLICANT: Tanzi, Rudolph E.			
APPLICANT: Kovacs, Dora			
APPLICANT: Saunders, Aleister J.			
TITLE OF INVENTION: Alpha-2-Macroglobulin Therapies and Drug Screening Methods for			
TITLE OF INVENTION: Alzheimer's Disease			
FILE REFERENCE: 0609.446003			
CURRENT APPLICATION NUMBER: US/09/241,606			
CURRENT FILING DATE: 1999-02-02			
NUMBER OF SEQ ID NOS: 27			
SOFTWARE: PatentIn Ver. 2.0			
SEQ ID NO 2			
LENGTH: 1474			
TYPE: PRT			
ORGANISM: Homo sapiens			
US-09-241-606-2			
Query Match	19.6%	Score 1441; DB 4; Length 1474;	
Best Local Similarity	29.1%	Pred. No. 5.9e-11;	
Matches 451; Conservative 266; Mismatches 551; Indels 284; Gaps 59;			
QY	12	LICVCTALAAVAGPRFLVTAAGIRPGCNVTVIGVLEHCPQVTVRAELTKTASNTLV	71
DB	15	LIVLPTDASVSGKQVAVVPSLHT-ETTEKGVLSYLNELTVTSASLESVAGNLSL	73
QY	72	-SVLEAG-VERKSGFKTLTPSLPLNSADE--LYELVGTGTODEILFNSNTRLSFET	126
DB	74	FTDLEAENDVHCVAF-----AVPKSSNBEVMEFLVQVKGPTQE---FKKRTVMVYN	124
QY	127	KRISVIFTDALKYKQKQKFRIVTFSDEPKRYTSLNLT-IXPKSNLTQOMLSQSD	185
DB	125	EDSLVFQVDTISYKPGQVTKFRVVSMDENRPLNELPLVLIQPKNRIAQWQSFLE	184
QY	186	LGVISKTRQLSSHPILIDGMSIOVQVND--QTYGSGFQVSEVYLPRFEVTLQTPILCSMNS	243
DB	185	GLKMGFSPLSSEPPGSGYKVVQKSGGRTEH-PTVBERVLPKFEVQVTVPKITITILE	243
QY	244	KHLNLTITAKTYTKKPKVGDVTLTFLPI-----SPWKG	276
DB	244	EMMNVSVCGLTYTGKPVGHVTVSICRYSDASDCHGDSQAFCFKFGQLNSHCFYQO	303
QY	277	KNKTKTKKINGSANFSFNDEMKKVMDSNGSLSEYLDLSPGPVILIT-----TVTSV	331
DB	304	VK--TKVFOILR-----KEYEMK-----LHTEAOIOEGTVLVTGSSSEITRTI	347
QY	332	TGISHNVSSTNVFQKHODYIIIEFPDYTVLTKPSLNFATATVKTADGNQLTLEERRNNVI	391

```

Db      348 TKLS-FVKVDHFRQG---IPFG-----QVRLVDGKGVPIP-----NKVI 384
Qy      392 TVTORNTTEWMSGNSGNQKMEAVQ-KINTY--VPGSTFKIEPP-----IIEDS 438
Db      385 FI-RGNENANTY--SNATTDHGLVQFSINTTNVMTSLTVRVNKKDRSPCYGOWVSEH 441
Qy      439 SELQKAYFLGSKSMVAHSLFKSPSKTYIQLKRDENIKVS-----PELVVSGNK--R 492
Db      442 BEAHHTY-----LVFSPSKSFVHLEPMSEHLPCHGTQYQAHYIILNGTLLG 489
Qy      493 LKELIS--YMWVSQQLVAVG-----KONST---MPSLPENSWTPKACVITYIEDDG 540
Db      490 LKKLSFYLLIMAKGIVRTGTHGLLVQEDMKGHFSISIPKSDIAPARLLIYAVLPTG 549
Qy      541 EIIIDVLKIPQVLKPKIKLYMSKTAEPSEKXSLASVQPSIYGIYAVDVSVLNM 600
Db      550 DVIDGSAKYDVENCLANKVDLSFSPQSLPASHAHLRYTA--POVCALRAVDQSVLLMK 608
Qy      601 ASNDITMENNVHEL-ELYNTGY-----YLGMPMS-----FA 631
Db      609 PDALSSASVYNLLPEKDLTFPPGLDQDDDCINRNVTYINGITTPVSTNEKDMYS 668
Qy      632 VFQECGLMVLTDANLTKDYIDGVYDNE-----YAEPMENEGHIYDIDHESL 680
Db      669 FLEDMGLKAFNTSKIRKQKCPQLOQYEMHGPBGLRVGFESDVWGRGHALVHVEE--- 725
Qy      681 GSSHH--VRKHFEPTWMLDTNMGRYLYQFEVTVDPDSTISWATGTVSIEDGLGLTT 737
Db      726 ---HTEVVRKYFPEPTWIMDLVVVNSAGVAVGTVPDITTEWKAFCFCEJSDGLDIS 782
Qy      738 TPVELQAFOPPEFLANLPSYVIRGEPEFALETITFNLYKDATEVAVYIEKSKDPILMTSS 797
Db      783 T-ASLRAPQPPFVELTMRPYSIRGEAFTLKATVLANLPKCIKRVSVQLEASAPLAVVEK 841
Qy      798 E-----INATGHQOTLLVPSSEDAVLFPIRPHLGEIPITVTLA-----PTA 841
Db      842 EQAPHCICANGRO-----TWSWAVPKSLGVNFTVSAEALESQELCTGEVPSV 890
Qy      842 -----SDAVTOMIIVKAEIGKYSQSILDLTNRLOSTIKTISFSPPTVTGSRVQ 896
Db      891 PEHRKQTVIKPLVPEEGLEKETTFNSLLCPSGEVS---BELSLKLPVNVESARAS 947
Qy      897 ITAGDVLGSPINGLASIRMPYCGGEONMNPANITYIDYLTKKKQOLTNLEKALSIF 956
Db      948 VSVUGDILGSMQNTQNLQMPYCGGEONMNPANITYIDYLTQQLTPEIKSKAIGY 1007
Qy      957 MRQGYRELLYQREDSFSFAG-NYDPS-GSTWLSAPVLRCLFLEADPYIDIQVNLHRTY 1014
Db      1008 LNTGYQQLANVKHYDGSYSTGERYGRNQNTWLTAFVLTKEFAQARAVIFIDEAHITQAL 1067
Qy      1015 TWLKGHQKSGNEFDPGRVHISELOGKNSPYTTATVTVSLGKRYKQPNIDQESIH 1074
Db      1068 IWLQORQKQNCFFRSGSLNNALKGVDEVTLSAITYTALLPLTVTHPPVRNALFC 1127
Qy      1075 LES-----EFSRIGSDNYTLALTLYALSIVGS-PKAKELANMLTWABQCGMFW--- 1124
Db      1128 LESAMKTAQEGDHG-SHYTTALATVAFALAGNDKREKVLKSLNEEVKXONSVMHERP 1186
Qy      1125 VSSSEKLSDSWQPR--SLDIEVAAYALSHFLQFQ---TSEGP---IMKWLRSQRNSL 1175
Db      1187 QKRPAPYGHFEYEPAPSAEVEMTSYVLLA-YLTNQPAFTSBDLSANINIVMITKQONAO 1245
Qy      1176 GGFASDTQDTVALKALSEFALMNTERTNIQVTVGSSPSPLAVQ----- 1222
Db      1246 GGFSTQDTVALALSKYGATFT-RTGKAQVYTIQSSGFSSKFPQDNNRRLILQOVS 1304
Qy      1223 ---PMAVNISANGFCAICQLANVYVYKASGSSRRRSTIONQAFDLDAVAK---ENKD 1275
Db      1305 LPELPGEYSKMYTEGCVYIQTSLKYNL-----LPEKEFPFAPAGVQTLPTQCD 1353
Qy      1276 DLN-LHVDLNVCTSPSGP---RSGMALMEVNLISGFVWPSBAISL---SETVKEVEYDHG 1328

```

```

Db      1354 EPRHNTSFOISLSVSYTSGRSASNMALVDYKWSGFIPLKPTVKLEERSNHVSRTVESSN 1413
Qy      1329 KNLVYDSVNETQFCVNIPAVRNFKYSNTODASVSYVDYEPERQAVRSYNS 1380
Db      1414 HVLITDKNSQTLSTLFTVYLQDVPVRDLKPAIVKYDIYETDEPAIAETNA 1465

RESULT 2
US-08-447-411-2
; Sequence 2, Application US/08447411
; Patent No. 5773243
; GENERAL INFORMATION:
; APPLICANT: FRITZINGER, DAVID C.
; APPLICANT: BREDEHORST, REINHARD
; APPLICANT: VOGEL, CARL-WILHELM
; TITLE OF INVENTION: DNA ENCODING COBRA C3, CVF1, AND CVF2
; NUMBER OF SEQUENCES: 81
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
; ADDRESSER: P.C.
; STREET: 1755 S. Jefferson Davis Highway, Suite 400
; CITY: Arlington
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/447,411
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/043,747
; FILING DATE: 07-APR-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Oblon, No. 5773243man F.
; REGISTRATION NUMBER: 24,618
; REFERENCE/DOCKET NUMBER: 1126-101-0
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 413-3000
; TELEFAX: (703) 413-2220
; TELEX: 248855 OPAT UR
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1651 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-447-411-2

Query Match      13.2%; Score 973.5; DB 1; Length 1651;
Best Local Similarity 23.5%; Pred. No. 9,8e-72;
Matches 391; Conservative 278; Mismatches 649; Indels 343; Gaps 61;

Qy      1 MGPPLITAAHLCTCTAALAVAPGRRELVTAGIIRPGANTVIGYEL-EHGPSQVTVK 59
Db      1 MEGALYLVALLIGPPGS---SHGALYTLITPAVIRTDTEQILVEAHGDSIPKSLDIF 57
Qy      60 A-----ELTKASHLTVSYLEABGVFKGSKTLPLSPLSNADSEIYELRVGTODEIL 115
Db      58 VHDPPRKQKTLFQSRVDMNQSGMFVT--PTIKVPAKELNKSQKQNYVAVKVGQVA 114
Qy      116 FSNSTRLSFEFKRISYFIQTDKALYKPKQEVKFRIVTLFSDFKPYKTSNLILKDPKSNL 175
Db      115 LKAVVLLSYSG--FFFIQTDKGIYTPGSPYRVKFSV--DHMHHRMDKTVIYE----- 164
Qy      176 IQQWLQOQSDLVYISKTFQLSHPI-----LQDKSIQOVQV---DOTTY 216
Db      165 -----FQTPREGIV-----VSSKPVNPSGSIIRPYNLPELVSGTWRKAVAKYHSPESY 213

```

QY 217 OSFOVSEYVLPKFEVTLQTP---LYCSNMSKALNGITTKYGVKQVDTLTLPLPSF 273  
 Db 214 AFEDREYVLPSEVTLQSPDKFLYIDGN-KNFHVSITARYLGVKVES-VAFWGVGYXI 271  
 QY 274 WCKKXNI---TKTEKINGSANFSPNDEMKVMSNGLSEYIDLSSGPEILLTWTYE 329  
 Db 272 DDAKSIIPDSLTRIPIDSDGEATLKRDLSRFPDNLQVHLLVYS-----VTVITE 325  
 QY 330 SVTGISRVNSTVNFPEKOHDIIEFPDYTVLKPISNFTATVYTRADGNQLTLEERRNV 389  
 Db 326 SGSDMVVTEGGGIHIVTSPYQYFRTKPKYFKGMPYELTVVTPDGSPPA-----H 378  
 QY 330 VTTTQKNTBYWSSGNSGNQMEAVOKINTVP---OSGTFKIEPPIEDSSELOKAY 446  
 Db 379 VPVVEALHSE-----GTLSDGTAKLINTPLNIQS-----LPITVRTHGDLPRE 425  
 QY 447 FLGSSSMVAHSLFKSPSKTYIQKTRDNIKVS--PELVVSGN---KALKELSTNV 500  
 Db 426 ROAITSMTATVATOGGSENYLHVAITSTEIKPGDNLVFNENVNGANSNLQIKFTYLI 485  
 QY 501 VSRGOLVAVGK-----ONSTMFSLTPENSMTPKACVIVYIIEDDGEIISDVLKIPVQ-- 552  
 Db 486 LNKGIIFKVGRRDQGLVTMNLHITPDLIPSRFAVYOVGNNEIADSVWVDVQDT 545  
 QY 553 ---LVFK---NKIKLYMSKVAEPSEKVSLSRISTVOPDSIVGIVAVDKSVNLMMAS 602  
 Db 546 CMGTLVVGKASRRDRIO-----KPGAAMKIKLE-GDPGARVGLVAVDKAVVYVYVNDK 596  
 QY 603 NDITMENVHELELNTGYLLGMFMSFVPOECGIMVLTIDANLT----- 647  
 Db 597 YKISQAKIWDTEKSDFGCTAGSGONNLGVFEDAGALATLSTNLTWKORSAKCPQANR 656  
 QY 648 ---KDYIDGVYDN---AEYAFPMENEG----- 670  
 Db 657 RRRSSVLLDSTAKAAQPODQGLAKCCGDMHNPMTGCKRAKTIQEGACRAAFIE 716  
 QY 671 -HIV-----DIHDFSLGSSPHV-RKHPEETIML-----DTWNG 702  
 Db 717 CCHYIKGIRDENQRESEFLARSDPEDELFGDDNISRSDFPESMLWLEELTGEBNNG 776  
 QY 703 YRIYQFEVTVP---DSTTSVWANGFVISEDLGLGTTT-----PYELOAFOPPFIF 751  
 Db 777 IS---SKTVPFYLRDSTITW-----ELLAVGLSPTKGIQVAAPEYIITVMDFID 823  
 QY 752 LNLPSVIRGEFEALETIFENYLDKATEYKVIIEKSDKEDIMTSSEINATHOQTILVP 811  
 Db 824 LRLPSVAVNBOVEIRAILYNTADEDIYRVELIYNPAF---CSASTEGORRQOPFK 879  
 QY 812 SEDGATVLPFIRPTHLG-EIPITVYALSPASDAVTOMILVKAGIEKSYQSILDLIT 869  
 Db 880 ALSSRAVPFVIPLRQGLHDEVIASVREGELASDGVKKLKVPPGERKNITITIELDPS 939  
 QY 870 DNRLOST--LKTLSGSPPNVTYTSERVOITAGD---VLGPSING--LASLIMPYGC 921  
 Db 940 VKGVGTOELTVIAKLDKVPDTEVEETISVLDGPAQIIENSIDGSLXHLITTPSGC 999  
 QY 922 GEONMINFAPNI---YIIDYLTKKQJLDNLKEKALSFRQCYQRELLYQREDGSFSAFG 978  
 Db 1000 GEONNITMTSPVIATYIDATGQENLGVDRTEIAKQIMTYAOQVYKKAHDHYAALT 1059  
 QY 979 NYDPGSGTWLSAFVLRCELEADPYI-DIDQVLAHRTYWL-KGHOKSNGEPMDPGRVHS 1036  
 Db 1060 N--RASSSWLTVAVVKVLAAGNWKDISHEITICGVKLLINRQOPDQVFENAPVHIG 1117  
 QY 1037 ELQGNKS---PVTLTAIVTSLGGRKYQPNIDVOESIHLESESRG-----I 1083  
 Db 1118 EMIGGKGKAEPRASLTAFTVALLSRSV-----CKEQINIIDSINKATDYLKKYEXL 1172  
 QY 1084 SNNYTLALITVALSVSGPKAEALNMLTWRAQEGQGFVWVSSEKLSDSQSPSLDE 1143  
 Db 1173 QRPYTTALTAVALA-----ADDRLDDRVLAAGSTRNRW-----EENNAATHNIE 1218  
 QY 1144 VAAVVALSHFLQFOTSEGIPIRWKLSRQNSLSGFGFASQDITVALKALSEFALMTER- 1202

Db 1219 GTSVALLLKMKKKAEBGVVVRMLIDOKYVGTYGTOATVWFOALAEVEIQPPTHOD 1278  
 QY 1203 TNOIVTGPSSPSP-----LAVQPMAY-----NISANGFALQOLNVVAVK 1247  
 Db 1279 LNLDSITKLPEREVDERSINDRNAVQARTETKLNEDPTVASASDGRATMTTLTVNAQ 1338  
 QY 1248 ASGSSRRRSRSLONCEAPFLDVAV-----KENKDLNHDVLANVCTSPSGRSGMALME 1300  
 Db 1339 L--REANVCNK--FLDVDSVENVELNLQAKGKKAALRLKICTRYLIGVDSMTMTIID 1392  
 QY 1301 VNLSEFNVPSBAI-----SLSEYVKKVEYDH-----GKLNLYDSVNETQ-FCVNIIPAV 1349  
 Db 1393 ISMLTGFPEDAEDELKRLSNGVDYRYSKFEIINNMAQKGTVAIYLDKVSHEDECHFKXH 1452  
 QY 1350 RNFKVSNTODASVSVVDYEPERRQAVRSYNSEVKLSSDCLSDVQCGRPCEBGASGSHH 1409  
 Db 1453 KHFEVGFLOPSGVKYSYNNLDEOCTKPYHDPDKGLINKICHGNICCAEBTCSLMOQ 1512  
 QY 1410 -----SSVIFIFCFKLL-----YEMEL 1426  
 Db 1513 KKIDQLRIQKACQNVVYVYKTKLRLIEKDGNDIYFMDV 1553

RESULT 3  
 US-08-793-126-1  
 ; Sequence 1, Application us/08793126  
 ; Patent No. 5849297  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Harrison, Richard Alexander  
 ; APPLICANT: Fairies, Charles Timothy  
 ; TITLE OF INVENTION: MODIFIED HUMAN C3 PROTEINS  
 ; NUMBER OF SEQUENCES: 2  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: HALE AND DORR LLP  
 ; STREET: 60 State Street  
 ; CITY: Boston  
 ; STATE: MA  
 ; COUNTRY: United States of America  
 ; ZIP: 02109  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: Patent in Release #1.0, Version #1.30  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/793,126  
 ; FILING DATE: 07-FEB-1997  
 ; CLASSIFICATION: 424  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Baker, Hollie L.  
 ; REGISTRATION NUMBER: 31,321  
 ; REFERENCE/DOCKET NUMBER: 102286,377  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (617) 526-6000  
 ; TELEFAX: (617) 526-5000  
 ; INFORMATION FOR SEQ ID NO. 1:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 1663 amino acids  
 ; TYPE: amino acid  
 ; STRANDEDNESS: single  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: protein  
 ; US-08-793-126-1  
 Query Match 12.7%; Score 931; DB 2; Length 1663;  
 Best Local Similarity 23.0%; Pred. No. 3,6e-68;  
 Matches 376; Conservative 286; Mismatches 615; Indels 356; Gaps 62;  
 QY 3 GPPITAAHLICVCTALAVAGPPEFLVTADGIR-----PGNNVTIGVELL 49  
 Db 6 GPSLL---LILLTFLPLALG-SPWYSIITBNILRESEETMVLGAHDAQGDVPTVTVH 60

```

QY 50 EH-----CPBQVTVKAEKLLKTASNLTVSV-LEAEVGEFKSGFETTLTPLSLNSADEIY 102
D 61 DPEKLLVLSEKSEITVLPATNMGNTVFTIIPANEFPSEKGRNFVYQA---TFGQV 117
QY 103 ELARTGTODEIIFSNSTRISFETKRISVFIQTOKALKYKQAEKFIIVLPSFKRYKT 162
D 118 EKVVLVSLQSGYL-----FIQTDKTYTTPGSTVLYRIFVTHNKLTVGR 161
QY 163 SLNLIIDPKSNLIQ--WLSQSDLGVIKSTFOLSHPIIDGMSIQVQVND---QTYQS 218
D 162 TWMNIENPESIPVKQDLSLQNOGLVPLSMIDPELVNMGOMKIRIAYENSPQOVSTE 221
QY 219 FOVEEYVLPKEEVLQ--TPLCYSMNSKHLNGITTAITYGKPYKQDVTLTF----- 268
D 222 FEVKEVYLPSEVIVPEPEKFFYYIYNEKGLVETITAFELVCKYKVEGAFVIFGIQDEQR 261
QY 269 -----LPLSWGKKKNTTKTFKINGSANFSPNDEBKNNWDSNGSLSEYLDLSSPG 319
D 282 ISLESKRIPIEDSGEVLRSKVLDDGVONPRAEDLVGKSLVSA----- 338
QY 320 PVELITVTSVTGSIKSNVSTNVFFKQHDYIIEFFDYTVLKPSLNTFATVYKTVRADGN- 378
D 329 -----TIIHSGSMQVAKERSGIPYVSPYQIHFTKPKYFKPMFPLDVFVTPNPGSP 383
QY 379 --QTLBERNNVITYTORNYTEWGSNGSNQKMEAVOKINYTVPOSGTFKLEPYLE 436
D 384 AYRPVAVOGEDYVQSLQDGVAKLSINTHPQK-----PLSTIVRKQKQELS 432
QY 437 DSSLOLKAFLGSKSMVAHSLFKSPKTYIQLKT-----RDENIKVSGPEFLVVSQ 490
D 433 EABD-----ARTVQALPYSTVSGNSNNYLSLVRTELPEETINVLMLMDRAHE 484
QY 491 KRLKELSYMVVSRGOLVAVGK-----QNSTMFSLTPENSWTPKACIVVY---IEDD 540
D 485 AKIYTYTYLLNKKRLKAGQVREPGDVLVPLSTTTPFIFRFLVANYTYLLIGASQR 544
QY 541 EISDVLKIPVQ-----LVFKNKIKLWYSKVKA-----PSEKVSRLISVTPDSIVGI 589
D 545 EVVADSVWVVDVQDSVCSLVVKS-----QSEDDQVPVQCGQMTLKIEGDHARV-L 595
QY 590 VAVDKSVNLMAASDITMENVHELELYNGYIYGMNNSFAYVQEGCLWLTDA----- 644
D 596 VAVDKGVFLNKKKLTQSKIMDVEKADIGCTPGSKDAGVFSDAGLTFSSSGQOTA 655
QY 645 -----NLTDYIDGVVDN-----AEXAERFME 666
D 656 QRAELQCPAPARRRRSVQLTERRMDKVKYKPKELRKCCEBGMKEMPMRSCQRTTIFIS 715
QY 667 ENL-----GHIVDIH-----DPSIGSSPHV--RKHPEPTWIM- 696
D 716 LGEACKVFLDCCNVITELRQHARASHLGLARNSLDEDIIAENIYSRSEFPESWLMV 775
QY 697 -----LDTNMGRIYOEFEVTVPDSITSVNATGFIYSEDGLGLTTPEVLOAFQ 746
D 776 EDLKEPPKNGISTKL-----NNIPLKDSITWEILLAVMSDKKGI-CVADPEVTVMQ 827
QY 747 PEFIFLPLPVSVIRGEEPALEITIENTLKDTEVKYIIE--KSKKPIILMTSSINATGH 804
D 828 DPFIDLPLPSVSVNEQVEIRAVLYNY-RONQELKAVELVLLHNAFCSLATTKR---RH 882
QY 805 QOTLIVSEDEGATVLPPIRP--THAGEIPITVTALSPASDAVQMTLVAKGIESYSQ 862
D 883 QOTITIPKSLSLSPYIYVPLKTLQLEVENAAUYHHFISGVRKSLKLVPEGRBMKV 942
QY 863 SILIDLTDNRL--OSTLKTSLSFSPENTV-----TSESR---VOITAGDVLPSPING- 910
D 943 AV-RTLDPERLGRGVOKE---DIPRADLSDQVDTSESTRILLQGPVQMTEDAVDAE 998
QY 911 --LASTLIMPYCGCGONMIFAPNYYIIDYLTKKQULDNLKER---ALSTMRQOYQBEL 966
D 999 RLKRLIYTPSGCGONMIGMTPIVIAHYIYDETEQWKEFGLEKQAGLELTKGYTQOLA 1058
QY 967 YQEDGSFSAFGNVDPSGTSWLSAFVLRCLFLEADPYIDIQNVLHRTYTWL-KHQQSG 1025

```

```

D 1059 FROPSAFAFAVVRAP--STWLTAAVYVVKFSLAVNLIAIDSOVLCAVXWMLILEKQPDG 1116
QY 1026 EFWDPGRVHSELOG-----NKSPTVLTAYIYVLSLGYRKYQONIDV-OESII----- 1073
D 1117 VFOEDAPVTHQEMTIGGLRNNNEKOMALTAFLVLSL-----QKADICEQVNSLGCSIT 1170
QY 1074 ----PSEFSRSGISDNVTLALITVALSSVSGSPKAKEALNMLTWRAEOBGMQFWVSSBS 1129
D 1171 KADDFLEANY--WMLQSGSYVALAGVLAQMGRLKGPLLKFLL-----7A 1214
QY 1130 KSDSQW---PRLSDIEVAAYALLSHFLPQTSSEGI-PIWRWLSKRONSLGCFASIQDPT 1185
D 1215 KDRNWEDEPKOLYNTAETSALLA-LLOLKQDFVPVPMWMEORYYGGGYGTOATF 1273
QY 1186 VALKALSEFAA-LMTERNTIOYTVGSPSPSLA-----VYQPMVANS 1229
D 1274 WYFOALAOYQKADPDDQELNDVSLQPSRSSKTYTHIRIMESASLLRSEETKENEGFTY 1333
QY 1230 ANGFGEAICQLNVVNVKASGSSRRRSIONQAPDLDAV-----KENKDLNHVDL 1282
D 1334 ABEKGQGLSVVTMTHAKKQDLTKN-----FLKVTIKAPETBERPDQAKXTMIL 1386
QY 1283 NVCTSSGPRSGMALMEVNLISGFVWSE-----AISLSEYVKYVY-----DHCKLNL 1332
D 1387 EICTRYRGQDQATMSIILDSIMMTGFAPDPTDILKQLANGVDRYISKYELDKAFSDRYTLII 1446
QY 1333 YLDSVNELO-FCNIRAVNFKVSNQDASVSLVDYEPERRQAVRSENVSE--KISSQD 1388
D 1447 YLDKVSHEBDDCLAFVHOIFVNELIOPQAVKAYAYNLBESCTRYHPEKEDKLNK-- 1504
QY 1389 LCSDVOGCRPCED 1401
D 1505 LCRCDAE 1516

```

## RESULT 4

```

US-09-132-271-1
; Sequence 1, Application US/09132271
; Patent No. 6221657
; GENERAL INFORMATION:
; APPLICANT: Harrison, Richard Alexander
; APPLICANT: Faries, Charles Timothy
; TITLE OF INVENTION: MODIFIED HUMAN C3 PROTEINS
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: HALE AND DORR LLP
; STREET: 60 State Street
; CITY: Boston
; STATE: MA
; COUNTRY: United States of America
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/132,271
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/793,126
; FILING DATE: 07-FEB-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Baker, Hollie L.
; REGISTRATION NUMBER: 31,321
; REFERENCE/DOCKET NUMBER: 102286,377
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 526-6000
; TELEFAX: (617) 526-5000
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:

```

LENGTH: 1663 amino acids  
 TYPE: amino acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein  
 US-09-132-271-1

Query Match 12.7%; Score 931; DB 3; Length 1663;  
 Best Local Similarity 23.0%; Pred. No. 3,6e-68;  
 Matches 376; Conservative 286; Mismatches 615; Indels 356; Gaps 62;

3 GEPILTAHLICVCTALAVAPRRVLTAPGIR-----PGAVTIGVELL 49  
 6 GSSL-----LILHLPLALG-SPWYSITTPNILLRESEETWLEAHDAQGVPLVTVH 60  
 50 EH-----CPQVYAKELTKTASNLTVSV-LEAEVFEKSGFKLLTSLPLNSADELY 102  
 61 DDPGKKLVISSEKTVLTTPATNMGNVTFITPANREKSEKGNKFTVQA---TGTQV 117  
 103 ELAVTGRTODELLFNSNTRLSFETKRISVFIOTDKALYKQEVKFRITVLFSDFKPYKT 162  
 118 EKVLVLSLQSGYL-----FIOTDKITYPGSTVLYKRITVNHKLIPGR 161  
 163 SLNLIKDKPKSNLIQO-WLSQSDLGVISKTQLSHPILGWSIQOVND---QTYGS 218  
 162 TWMVNIENPEGIPVKQDSLSQNLGVLPISWDIPELVNMQWKIRAYENSPOQVFTSE 221  
 219 PVSSEVYLPKEFVTLQ--TFLYCSMNSKHLNGTITAKTYGKPVAGDVTLT----- 268  
 222 FEVKSEVLPSEFVIEVEKFIYIYNEKLEVTITARFLYKQVETARFVIGIDGEGR 281  
 269 -----LPLSPFGKKKNTTKFKINGSANFSFNDSEKQVMSNGLSEYLDLSPG 319  
 282 ISLPSLKRIPLEDGSEVVLRSKVLVDGVQNPRAEDLVGSLYSA----- 328  
 320 PVEILITTESVTGISRNVSTNVFKQHDYIIEFDYTVLKPISLNFATVKTADGN- 378  
 329 -----TVLIHSGSDMWQARSGIPVITSFYQIHFTKPKFKPMFDMVFTVMDGSP 383  
 379 -QTLERRNNVITVQANTETKSSNSGNQKAEVOKINVTYPOSTKIEPILE 436  
 384 AYRPAVAVQGEDVQSLTQGDVAKISINTHPSOK-----PLSITRTKQELS 432  
 437 DSEELQKAYFLGSKSMVAHSLFKSPSKTYIOLKT-----RDENIKGSPFELVSGN 490  
 433 EAEQ-----ATRMQALPYSTGNSNNYHLSTRLPGETLVNPLIRMDRAE 484  
 491 KKLKELSTVNVSRGOLVAVGK-----QNSTWESLTPENSMTPKACVIVY---IEDDG 540  
 485 AKIRYTYVLIIMKGRLLKAGQVREPGDVLVLPISITTDPIPSFLVAVYTLIGASGR 544  
 541 ELISVCLKIPVQ-----LVFKKIKLKYKVAE-----PSEKSLRISTVQPDISTYGI 589  
 545 EVVAASVWVDVDSQVSGSLVAVKSG-----QSEDRQVPCQOMTLKIEGDHGARV-L 595  
 590 VAVDSVNLMAANSNDITMENVHELELVNTGYVYLGFMFMSFAVPOCGMLVTLDA----- 644  
 596 VAVDGVAVLANKKNTLQSKIMDVVEKADIGCTPSGKQYAGVFPAGLFTTSSGGQOTA 655  
 645 -----NLTKOYIDVYND-----AEYAEFME 666  
 656 QRAELQCPARARRRSVQLTEKMDKVGKPKELKCCEDGMRNPMFSCQRTRTFS 715  
 667 ENE-----GHIVDIH-----DPSLSSPHV--RKHPEFTWIM-- 696  
 716 LGEACKVFLDCCNVITELRQHARASHLGLARSNLDEDIIEENIVSRSEPESEMLNV 775  
 697 -----LPTNNGYRITYGEFVYVPSITSWATGVISEDGLGTTTPVELQAFQ 746  
 776 EDLKEPPKNGISTKL-----MNIPLKDSITTWELAVSMDKGI-CVADPFETVWQ 827  
 747 PFIPLNLPYSVIRGEFALETITFNVLADATEVKVIL--KSDKFDILMTSSEINATGH 804

828 DFDILRLPYSVVRNEQVEIRAVLVNY-RQNEKLRVELLHNPACSLATTKR-----RH 882  
 805 QOTLLVPSDEGATVLPFRP--THLGEIPITVTALSPVADVMTQMLVKAIGIEKYSQ 862  
 883 QOTITIPKSSLSVYVVLPLKTGLQEVKRAVYHHHTISQVRSKLKVPBGIMNKTV 942  
 863 SLIDLDTNRL--OSTLKLTSFSPPTNV-----TGSER--VOITAGVLAGPSING- 910  
 943 AV-RTLDERLGRREGVQEE--DIPPALSDQVPPTSETRLLIGCTPVACQWTEADVADE 998  
 911 -LASLRMPYGGGEGENMFAPNITYILDYLRKKOULTNLK--AISFNGQIQRELL 966  
 999 RLKHLIVPSSGGEONMGMTPTVAVHYLBETEEMERFGLEROGALELIKKGTYOOLA 1058  
 967 YOREDSFSARGNDVPSGSTWLSAPVLCPLFADPYIDQNVHRTTWL-KGHOKNSG 1025  
 1059 FRQSSAPRAAFVYKRAV-STWLTAVVAVVFSLVANLAIIDQVLCGAVKWLILEKQPDG 1116  
 1026 EFWDPGRVHSELQGG-----NKSPTLTAVYVTSLLGYRKYQPNIDV-QESI- 1073  
 1117 VQEDAPVTHQEMIGGLANNKEMALTAFLVLS-----QEAKDICEQVNSLPSGIT 1170  
 1074 -----FLESEFSRGSIDNTTALITYALSSVSPKAKELNMLTWRAEQEGGQVNVSSS 1129  
 1171 KAGDELANY-MNLQRYSTVAIAGYALAQMGRLLKGPLINKFLT-----TA 1214  
 1130 KLSDSWQ--PRSLDIEVAAYALLSHFOFQTSSEI-PIMWLSQRNSLGGFASTOFT 1185  
 1215 KDKNWEQPGKQLVNVEATSYALLA-LIQLDKDFVPPVAVWMLNQRYYGGYGSTQATF 1273  
 1186 VALKALISEFAA-LMTERKNIQVTVGPSSPSPLA-----VQOPANVNS 1229  
 1274 MVFQALAQYQADAPHOGLNLDVSLQPSRBSKITHRIHWSASILRSEETENGFTYT 1333  
 1220 ANGFQFALCOLNVVNVKASGSSRRRSIQNEAFDLVAV-----KENKDLNHYDL 1282  
 1334 AEGKQGTSLVTVTHAKAKQLTCNK-----FDLKVTIKPAPETEKRPQDAKNTML 1386  
 1283 NVCTSFSGPGRSGMALMENVNLISGRVPS-----AISLSEVYKVEY-----DHGKNTL 1332  
 1387 EICTRYRDQDQATWSILDISMTGAPAPDTDLKOLANCVDRYISRYELDKAFSDNNTLIT 1446  
 1333 YLDSVNETO-PCVNI.PAVANFVSNTPDASVSIYDYEPFRQAVSYNSEV---KLSGCD 1388  
 1447 YLDKYSHSEDDCLAKVHQYRNVELDQGAKVAVAYVNLSESCRTFYHPEKEDGKLNK-- 1504  
 1389 LQSDVQGRPCED 1401  
 1505 LCRD-ELCRCAE 1516

RESULT 5  
 US-09-142-334-22  
 Sequence 22, Application US/09142334  
 Patent No. 6268485

GENERAL INFORMATION:  
 APPLICANT: Fairies, Timothy C.  
 APPLICANT: Harrison, Richard A.  
 TITLE OF INVENTION: Down-Regulation Resistant C3 Convertase  
 FILE REFERENCE: 4-30443/A/IMG/PCF  
 CURRENT APPLICATION NUMBER: US/09/142,334  
 CURRENT FILING DATE: 1999-04-15  
 EARLIER APPLICATION NUMBER: PCT/GB97/00603  
 EARLIER FILING DATE: 1997-03-04  
 NUMBER OF SEQ ID NOS: 35  
 SOFTWARE: Patent In Ver. 2.0  
 SEQ ID NO 22

LENGTH: 1663  
 TYPE: PRT  
 ORGANISM: Homo sapiens  
 US-09-142-334-22  
 Query Match 12.7%; Score 931; DB 3; Length 1663;

Best Local Similarity 23.0%; Pred. No. 3.6e-68;  
Matches 376; Conservative 286; Mismatches 615; Indels 356; Gaps 62;

```

Qy 3 GPRLLTAHLCVCTAALAVAPRPLVTARGIIR-----PGNAVITIGELL 49
Db 6 GPRSL-----LILLTHPLALG-SPWYSLITTNLLESEETMVLBAHDAQDVPVTVH 60
Qy 50 EH-----CPSQVTVKLELKTASNLTVSV-LEAGVPEKSGFKTLTLPILNSADEIY 102
Db 61 DFGPKVLISEKTVLTPATNMGVNTFTIIPANREFSEKGRNFVYQA---TFGTVV 117
Qy 103 ELRYTGRTQDEILFSNSTRLSFETKRLSVFIQTDKALYKPKQEVFRIYVLFSDFKPKYT 162
Db 118 EKVVLVSQSGYL-----FIQTDKTYTPGSTVLYRIFVYVNHLLFVGR 161
Qy 163 SLNLLIDPKSNLLIQ--WLSQSDLGVSITKTFQSLSHILGDMGSIQVQVND---QTYQS 218
Db 162 TVWVNIIEPBGIPYKQDLSQNOGLVPLSMWIPELVNMQMKIRAYENSPOQVSTE 221
Qy 219 FQVSEYVLPKFEVTLQ--TPLYCSMNSKHLNGITTAQYTKPKYKGDVTLTF----- 268
Db 222 FEVEEYVLPSEFVIVEPEKFPYIYNEKGLVITITARLYGKXEGTAFVIFGIQDGBQR 281
Qy 269 -----LPLSFMGKKKNTTKTFKINGSANFSFNDEKKNVMDSSNGLSSEYIDLSSPG 319
Db 282 ISLPESLKRIPIEDSGGEVLSRKVLLDGVQNPRAEDLVGKSLYSA----- 328
Qy 320 PVELITVTESVTVGISRVSTNVFFKQHDYIIEFDDTYVLKPSLNPATYKVRADGN- 378
Db 329 -----TYLHSGSDMVQAEKRSIGPIVTSFYQIHFTKPKYKPMPLMVFVTPGSP 383
Qy 379 --QTLBERRNNAVITVQRYNTEYSGNSGNOKMEAVOKINATVPOSGTFKLEPIL 436
Db 384 AYRPVAVQSGEDTYQSLQSGVAKLSINTHPQK-----PLSTIVRKQKELS 432
Qy 437 DSSELQKAYVLGSKSSMAVHSLFKSPSKTYIQKT-----RDENIKVSGPELVVSGN 490
Db 433 EAEQ-----ARTMQALPYSTVSGNSNNYHLHSLVLTETLRPGETLVNVLAMDRAHE 484
Qy 491 KRLKELSGVNVSRQOLAVGK-----ONSTMFGLTPENSWTPKACVIVY---IEDDG 540
Db 485 AKIRYTYLLNMKRLKAGQVRBPQDVLVLPSTTDFIPSRFLVAYTLLIGASQOR 544
Qy 541 EILSDVLKIPVQ-----LVFNKIKIKLYWSKYKAE-----PSEKYSRLISTVQPSIVGI 589
Db 545 EVVADSVWVVDYKDCSGSLVYKSG-----QSEDRQPVGQOMTKIBEDHARVY-L 595
Qy 590 VAVDKSVNLMAASNDITMENVHLELYNTGYILGEMFNNSPFAVQEGGLWVLTDA----- 644
Db 596 VAVDKGVFLNMKKNKLTQSKIWDVEKADICTPGSGKDVGVSFSDAGLFTSSGGQTA 655
Qy 645 -----NLTKDYIDGVYDN-----AEVAERPME 666
Db 656 QRAELQCPAPARRRRSVQLTKEMDKVQKPKELKCCCEGBREMPRESCQRTTFIS 715
Qy 667 ENB-----GHIVDIH-----PESLGGSPHV--RKHFPEPTWII-- 696
Db 716 LGECCKVFLDCCNYITELRRQBARASHLGLARSNLDEDIJAENIYSRKSFPESSMLNV 775
Qy 697 -----LDTNMGRVRIQEFVTVPDSITSWVATGTVISDILGLTTPVELQAFQ 746
Db 776 EDLKEPPNGISITL-----NMIPKDSITTWELLAVNSDKKI-CVADPREVYVWQ 827
Qy 747 PFFFLNLPYVINGEFALEITTFNYIKATERYVYIIE--KSKPDLMTSSINATGH 804
Db 828 DFFIDRLPYSVVNEQVEIRAVYVNY--RQNEQLKRVLELHNPFCSLATTKR-----RH 882
Qy 805 QQTLLVSEDCATVLPFIRP--THLGEIPITVTLSPASDAVYOMLIVYAEIGESYSQ 862
Db 883 QQTITTPKSLSPYIVLPLKTLQGEVCAVAAVYHHFISGVKSLKVVPEGIRANKTV 942
Qy 863 SILDLTDLNRL--OSTLKTLSFSPPTNV-----TGSER---VOITAGVVLGSPSING- 910

```

```

Db 943 AV-RTLDPERLGBEGVQK---DIPPADSDQVDPTESETRILLQGTTPVAQMTEDAVDAE 998
Qy 911 -LASLITMPYGCSEQNMNINFAPNIYLLDYLTKKKQOLDNLKEK---ALSPRQGYORELL 966
Db 999 RLKHLIYTPSGCCEQNMIGMTPTVIAVHIYIDTEQNEKRELEBROGALBKIKGYOQLA 1058
Qy 967 YQREDSFSAFGVNDPSGTSWLSAFVLRCELEADPYIIDQNVLHRTYWL-KGHQKNSG 1025
Db 1059 FQPPSAFAFAFVKRAP--STWLTAAYVVKVPSLAVNLIAIDSYLCAVKKVLLIEKQPDG 1116
Qy 1026 EFWDPGRVHISELQGG---NKSPTVLTAYISVLSGKRYQGNIDY-QGSIH----- 1073
Db 1117 VFQEDAPVHQEMIGGIRNNNEKDMALTAFLVLSL-----QAKDICEQVNSLPGSIT 1170
Qy 1074 -----FLESEFSGISDNYTLALITVYLSVSGSPKAEALNMLTPRAEQEGMQFWVSES 1129
Db 1171 KAGDFLEANY-NMLQNSYVIAIGYALAOGRKLGPLINKFLT-----TA 1214
Qy 1130 KLSDSWQ---PRSLDIEVAAYALLSHLOPQTSBGI-PIWRMLSRONSLGFASTQDPT 1185
Db 1215 KDKRMEDPKQOLYNVATSYALLA-LLOQKDFDPVPVVRMLNEORYYGGYGTQATF 1273
Qy 1186 VALKALSEPA--LMTERTVIOYTVTGSPSPFLA-----VQPMVAVNIS 1229
Db 1274 MVFOALAQYOKADPDHQLNLDVSLQPSRSKITHRIHESASLSRSEETKENEGFTVT 1333
Qy 1230 ANGEFPAICQNAVYVNVKASGSSRRRSIQNEAFDLDVAN-----KENKDDLHVDL 1282
Db 1334 AEGKQGTLSVMTAKKADQLTCKN-----FDUKITIKAPETEKRPQAKMTML 1386
Qy 1283 NVCTSPSGRSQMALEVNLLSGFVPSB-----AISETVKKEY-----DHGKML 1332
Db 1387 EICTRYRGDDAMTSLIDISMGTGAPADPTDQLQANGVDRYISKYELDKAFSDRATLLI 1446
Qy 1333 YLDSVNETQ-FCVNIIPAVNFKVSNQODASVSLVDYIEPRQAVRSYNSG---KLSQCD 1388
Db 1447 YLDRKVSSEDDCLAFVHQYFVNLIQPGAVKYAYVYVLEESCTRFYHPEKEDGKLNK-- 1504
Qy 1389 LCGSDVQCGRCPED 1401
Db 1505 LCRD-ELCRCAEE 1516

```

```

RESULT 6
US-08-487-283A-2
: Sequence 2, Application US/08487283A
: Patent No. 635245
: GENERAL INFORMATION:
: APPLICANT: Evans, Mark J.
: APPLICANT: Matlis, Louis A.
: APPLICANT: Mueller, Eileen Elliott
: APPLICANT: Nye, Steven H.
: APPLICANT: Rollins, Scott
: APPLICANT: Rother, Russell P.
: APPLICANT: Springhorn, Jeremy P.
: APPLICANT: Squinto, Stephen P.
: APPLICANT: Thomas, Thomas C.
: APPLICANT: Wilkins, James A.
: TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE TREATMENT
: TITLE OF INVENTION: OF INFLAMMATORY DISEASES
: NUMBER OF SEQUENCES: 26
: CORRESPONDENCE ADDRESS:
: ADDRESS: Seth A. Fidel
: STREET: 25 Science Park (Alexion)
: CITY: New Haven
: STATE: Connecticut
: COUNTRY: USA
: ZIP: 06511
: COMPUTER READABLE FORM:
: MEDIUM TYPE: 3.5 inch, 1.4mb storage
: COMPUTER: Macintosh Cetrus 610
: OPERATING SYSTEM: System 7
: SOFTWARE: WordPerfect 3.0

```



CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/487,283A  
FILING DATE: June 7, 1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/236,208  
FILING DATE: 02-MAY-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Seth A. Fidel.  
REGISTRATION NUMBER: 38,449  
REFERENCE/DOCKET NUMBER: ALX-152.1 CIP  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (203)776-1790  
TELEFAX: (203)772-3655  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1676 Amino Acids  
TYPE: Amino Acid  
STRANDEDNESS: Single  
TOPOLOGY: Linear  
MOLECULE TYPE:  
DESCRIPTION: Pro-C5 Polypeptide  
HYPOTHEICAL: No  
ANTI-SENSE: No  
ORIGINAL SOURCE:  
ORGANISM: Homo sapiens  
PUBLICATION INFORMATION:  
AUTHORS: Haviland, D.L.  
AUTHORS: Haviland, J.C.  
AUTHORS: Fleischer, D.T.  
AUTHORS: Hunt, A.  
AUTHORS: Wetzel, R.A.  
TITLE: Complete cDNA Sequence of Human  
Patent No. 6355245  
JOURNAL: Journal of Immunology  
VOLUME: 146  
PAGES: 362-368  
DATE: 1991  
US-08-487-283A-2

Query Match 12.5%; Score 921; DB 4; Length 1676;  
Best Local Similarity 24.0%; Pred. No. 2.5e-67;  
Matches 379; Conservative 297; Mismatches 621; Indels 284; Gaps 67;

28 FVTAAGIIRPGNVTIGVEL---LEHCSQVTVKAEELKTSANLTVSY-LEAEQFEKG 83  
23 YVISAQKIFRGASENVIVQVGYTEAPATISIKSYDPKKFSYSGHLSSENKFNQNS 82  
84 SKRTLTLPSP--LNSADEIYELRVGRYODEILFNSNSTRLSFETGRISVFIOTDKALYK 141  
83 ALITTPQQLPGQGNVSVYV-LEVVSKEH-----FESKGRMPTTYDNGFLFHTDKPYVT 136  
142 PQGEVAFRIVTLPSDFPYKTSINLIIKPKSNLIQQWLSQOSDLGVIS-KTFQULSHPI 200  
137 PQOSVAVRVYSINDDLKPAKRETVLTFIDPEGSEVDM-VEBIDHIGIISFPDPFKIPSNR 195  
201 LGDMGIVQVGNQ-----TYOSFQVSEVVLKFEVTLQTPLYCSN---NSGHLNGTIRA 252  
196 YGMWITKAKYKEDFSTGTAY--FEVKEVLLPHFVSIE-PEVNFPGYKNFNFETITKA 252  
253 KTYGKPV-KGDTVLTFLPLSPFGKKKN-----TKTFKINGSANFSFNDEEM 299  
253 RFYKRVVTEADVYILF-----GIREDLKDOQKEMQOTAMOTMLINKIAQVTPDSE-- 304  
300 KAVMDSNGLSEYL--DLSSPGBEVILTTVESVTGISNVS--TNVFKQHDYIIIEFDY 356  
305 ---TAVKELSYYSLEDLNNK--YLVIATVVISSTGSGFSEAEIPGIKYVLSFYKLNIVAT 359  
357 TVVLKSLNPTATVTKVTRADGNQLTLEERNVNI-----TYQRYVTEYMGSGNSGNK 411  
360 PLFLKRGIVPIKVVQ-----KDSLDQVGVGPVVLNACTIDVNOETISDLDPSKSVTVK 413  
412 MEAVQKINTVPOSQTFKIEPILEDSESLQKAYFLGSKSSMAVHSLFKSPSKTYIQLK 471

414 DDGVASFYVNLNP-SGVTALFEFNVKTDADPLEENQAREGRALAVSSL--SQSYLIYDWT 470  
472 TRDENIKYGSPEELVVSQNK-----RLKELSYMVVSRGQVAVGNQNS-----TMFSILTP 521  
471 DNHKALLVGEHLNITVTPKSPYIDKITHYNNLISKGIHFGREKFSQDASYSINIPV 530  
522 ENSWTPKACVIVVYI---EDDEIISDVLIKIPVOLFNGIKIYWS--KVRAEPEKYSL 576  
531 TQNMVPSRLVYIYVNGEQRAELVSDSVMLNIEKGNQCVHLSPPADAYSPOQYSL 590  
577 RISTQPDISIYIYVAVDKSVNLMNASNDITMENVVEHELYNTGYIIGMFNNSFAVPOEC 636  
591 NMA-TGMDSVWALAADVDAVVGQAKKPLERVQPLEKSDLGCGGAGGJNNANVFLHA 649  
637 GLMVLTDAN-----LTKDYIDGY-----YDNA----- 658  
650 GLTFLTNANADDSOENDEPCKEILRPRTLOKKIEELAAKYHSHVKKCCYDGAQVNNDE 709  
659 ---EYAEERF-----MEENEGHVIDHDSIG-----SSPHV 686  
710 TCEQRAALISLGRICAKFTECCVVASQLRANISH---KMOGLRLMKTLPLVSKPEI 765  
687 RKHPETWIMLDTMNGYRIYQEFVTVBDSITSVWATGFI SEDLGGLATTPVELQAFQ 746  
766 RSYFPESWLM-BVHLVPR-RKQGLAPDLSLTWEIQIGIS-NTGICVADT-VKAKYFK 821  
747 PFIITLNPYSYVIRGEFALETTFNLIKADTEKVYIIEKDKPDILMTSEINATGQQ 806  
822 DVLEEMNIPYSVVRGEQIQLKGYVNYRTSGMFCVKNASAVEG---ICTSESPYIDHQ 877  
807 T-----LWVSEDAVTLFPIRPTHGELPIYVTLASPTASDAVTOMLIVKAGIKS 859  
878 TKSSKCAQKQKVBGSSSHLYFTVPLPLEIGLNINPSETWGEKELVKTIRVPEGVARE 937  
860 YSOSIILDLTNRLOSTL---KLSFSPPTVTVGSEKVOITA-----IGDVLGP--SIN 909  
938 SYSQVTLDP--PRGIYGISRRKEPPYRIPLDLVPKTELKRLTSYKGLVGEILSAVLSQE 995  
910 GLASLIRNPGCGEGNMIFAPNIIYIDY-----LTKKQLTDLNLEKALS 955  
996 GINILTHLPKSAEELMSVVPVFPVHYLETGNHMNIHFSDPLIEKQKLKKKKEGMLS 1055  
956 FMQGYQGEILLYQERDSEFSAFGNVPDSGNTWLSAFVRLCGLADPYDIDONVLRHYT 1015  
1056 IMS-----YKRNADISYVWKG--GSASTWLPALFALRVLGQVNVKEQNSICNSLL 1105  
1016 WL-KGHQKNGEFMDPGRVHISELQ-----GNKSPVLTAVIYVTSLLGYRK---YQPN 1066  
1106 WLVENYQLDNGSPKENSQYQIKQGTLPVARENSLYLTAFVY---IGIRKAPDICGLV 1162  
1067 DVQESI---HFLSEFSGKISDNYTALITVALISVGS---PAKKAELNLTWRABOEG 1119  
1163 KIDTALIKRADFL-LENTLPAQSTFTLLAISAYL-SLDGKTHPQFRSIVSALKEBALYK 1220  
1120 G---MQFVSS-----ESLSDSMQPRSLDIEVAAYALLSHLQCTQTSIGIIMWLSQR 1172  
1221 NPPIYRFKMDLQHKDSSVPNTGTARM--VETVALALTSINLKQIDINYNVPIKMLSEBQ 1278  
1173 NSLGFASSTQDPTVVALKALISEFALMTERTNIQVTVGPSSPS-----PL 1218  
1279 RYGGGFYSTQDTINABELTSEYSLVQKRLSMIDIVSYKEKGLAHNYKMTDKNPLGKPV 1338  
1219 AVQPMAVN---ISANGGFALCOLNVVYNYKASGSSRRRSIQNQ-EAPDLVAVKENK 1274  
1339 EYL-----LNDLLIVSTGSGGLATVHTVHAKTISTSEVCSFYLIKIDTODIEASHYRGY 1394  
1275 DDLNHNVDLNVCTSGSGPR-----SGMALMEVNLISGFMPSEAL-SLSTFYKVEEDH 1327  
1395 GNSDTKRIVACASVK-PREBSSGSSHRAVNDISLPTGISANEBDLKALVGVQDLFPDY 1453  
1328 ---GKLNLYLDSVNETQF-CVNIPAVNFVSNTOVASIVDVYEFRRQAVNSYNSEV 1382

Db 1454 QIKGHHVILQJNSIPSSDFLCRRRRIEFLFEVGLSPATFTVYVHRHPRKQCTMFYST-- 1511

Qy 1383 KLSGCDLSDVQGRPCEDGA 1403

Db 1512 -----SNIKIKVCEGGA 1524

RESULT 7

US-08-662-227-2

Sequence 2, Application US/08662227

Patent No. 592320

GENERAL INFORMATION:

APPLICANT: VOGEL, CARL-WILHELM

APPLICANT: BREDEHORST, REINHORST

APPLICANT: KOCK, MICHAEL

APPLICANT: FRITZINGER, DAVID

TITLE OF INVENTION: RECOMBINANT PROCVF

NUMBER OF SEQUENCES: 39

CORRESPONDENCE ADDRESS:

ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,

ADDRESSEE: P.C.

STREET: 1755 S. JEFFERSON DAVIS HIGHWAY

CITY: ARLINGTON

STATE: VA

COUNTRY: USA

ZIP: 22202

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: IBM PC compatible

SOFTWARE: Patent in Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/662,227

FILING DATE: 14-JUN-1996

CLASSIFICATION: 530

ATTORNEY/AGENT INFORMATION:

NAME: OBLON, NORMAN F.

REGISTRATION NUMBER: 24,618

REFERENCE/DOCKET NUMBER: 1126-0107-0X

TELECOMMUNICATION INFORMATION:

TELEPHONE: 703-413-3000

TELEFAX: 703-413-2220

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 1642 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-662-227-2

Query Match 12.2% Score 894; DB 2; Length 1642;

Best Local Similarity 22.5% Pred. No. 4.4e-65;

Matches 361; Conservative 295; Mismatches 672; Indels 276; Gaps 56;

Qy 18 AALAV-----APGRFLVTAFGIIRPGNVTIGVELL-EHCPSOVTVKA-----ELKKTAS 67

Db 10 AALAIIGPSSHGALYTLITPAVLTITREBOILVEAGDSTPKDDITFVHDFPKQKTLF 69

Qy 68 NLTVSVLEAGVFEKSGFKTLTLPSPILNSADEIYELRVGTGRTODEILFSNSTLSPEFK 127

Db 70 QTRVDMNPAGMLVT---PTIEIPAKEVSTDSRONQVYVVOVTPQVRLKLVLSYQSS 126

Qy 128 RISVFIQTDKALYKQDEVKFRITVLTBSDFKPYTSLNLIKDKSKNLIQMLSQGSDLG 187

Db 127 --FLFIQTDKGIYTPGSPVLRYFVSSMDNTSKANKTYIVEFQTEGILVS---SNSVDLN 181

Qy 188 VISKTPOLSHRILGDNSI--QVOVNDQTYQSFQVSEVYLKPEPVTLQTP---LYCSMN 242

Db 182 PF-WPYNLPDLVLSGTRIVAKYEHSPENYATYVDVKRYVLPSEVRLQSEPKFFYIDGN 240

Qy 243 SKHLNGITTAQYKYGVKPYKGVTLTFLPLSPFWGKKKI---TKTFKINGSANFSFDEE 298

Db 241 -ENFHVISTARVLYGEEVSG-VAFLFQVKIDDAKSIPOSLTRIPILIGDGKATLKRD 298

Qy 299 MKAVMDSNGLSEYLDLSSPGVEILITVYESVTGISRVNVTVPFQKHDIIEFPDYTT 358

Db 299 FRSRFNLNLELVCHTILYAS-----VTWTESSGDVVVTEQSGIHLVASPYQHLFKTPK 352

Qy 359 VLKPSLNFTATVAVTRADGNQLTLERNNNVITVTQRNYTEWSSGNSGNGMEAVQKI 418

Db 353 YFPRGMPVEYELTVVTPDGSFA-----AHVPVSEAF-----HSMGTTLSDTAKL 398

Qy 419 NTVTPQSGTFKIEFLLEDSSQLKAYFLGSSKSMVSHLSPKSPKTYIQLTRDENTK 478

Db 399 ILNIPILNAQ---SLPTVTRNHDDLPREROATKSMATLAVOTQGGSGNVLHVAITSTEIK 455

Qy 479 VGS--PFELVYVSN-----KRLKELSYVAVSRGLVAVGK-----QNSTVFSITPENSWT 526

Db 456 PGDNLVNFVFNKGNASLKOIKFTYLLINKGKIFVNGQPRRDGNVLTNMLHITPDL 515

Qy 527 PKACVIVYIYEDDEIISDVKLPVO-----LVFR--NKIKLYSKVKAEPSEKVSRLI 578

Db 516 PSFRFVAYYQVGNNEIIVADSVWVDVQTCGTLVVGKDNLIQ-----PGAAKKIKL 567

Qy 579 SVTPQDSIVGIVADSKVNLMAANDITMENVHELIVNTGYLGMFMSFAVPOEGL 638

Db 568 E-GDPGARVGLVADRAVYVLDKRYKISQAKIWDTEKSDFGCTAGSGQNNLGFEDAGL 626

Qy 639 WVLTDNL-TRD-----YIDGYD-NAEYAEFRM--EENE-GHIYDIHF 678

Db 627 ALTTSTNLTNKKSAKCPQANRRRSVLLDSNASKAERQDDDLAKCCEDVNHENP 666

Qy 650 -----YIDGYD-NAEYAEFRM--EENE-GHIYDIHF 678

Db 687 MGYTCERAKYIQGDACKNAFLCCRYIKGVADENQRESEFLARDNDEGFIADSDII 746

Qy 679 SLSSPHVAKHPEPTIWL-----DTMGRYLYQEFYVTPDSITSWATGVISEDG 732

Db 747 S-----RSDPFKSWMLTKDLTEEPNQGSSKMSFYRLSITVWVLAVSFPPTKG 799

Qy 733 LGLTTTPVELQAPQPFIFINLPYSVIRGSEFALETTFINYLDATEVVKYIEKSDFDI 792

Db 800 I-CVAREYELRWKVFIDLOMPYSVAKQEQEIRAILNRYVVEDIYVVELLYNPAF-- 856

Qy 793 LMTSSEINATHQQTLLVPSDEGATVLPPIRPTHLC--EIPITVVALSPASDAVTOMIL 850

Db 857 --CSASTKGQRYQCPPIKALSRAVPFVYLPLEGILHDEIKASVQELMSDGVKKLK 914

Qy 851 YKAEGLKESYSGSILLDLTDNRLOST-----LKTLSFSPRPVTYVSGSERVOITAGDVL 904

Db 915 VPEGVQKSTVITVKLDPPRAKGVGTQLEVIKARKLDRAVPDTEIEFTKIIIQGDPVAQII 974

Qy 905 GPSING--IASLIRMPYGCQGMNINFAPNI---YIIDVLYTKKKQLTDLNLEKALSPMRQ 959

Db 975 ENSIDSKNLHLITTSQCGEQMIMMAAPVLTATYLLDTTEQWETIGINRRTEAVNOIYT 1034

Qy 1035 GYVQGMVYKKADSYAFTN--RASSSWLTAVYVKKFAAAKAVAGASIEIICGVRWLL 1092

Db 1018 KGHQKNGEFPWDPGRVYHSELQG--GNKSPVTLTAYITSLSIGR---KIQPNID--V 1068

Qy 1093 LNKQCPDGAFKENAPVLSGMOGIGAGAEVYLTAFIIVALLSESTKICNDVYNLSDSGI 1152

Db 1069 QESIHFLSEFSFGISDNTYTLALITVALSSVSPKKEALNMLTWABOEGGQMFVSS 1128

Qy 1153 KKAATNVLKTYER-LQRPYTLALTAYALA-----AADQLNDRVYMAASTGDH-- 1201

Db 1129 SKLSDSQPSLDIEAAYALASHFLQFQTSRPIRMWLSRQNSIGGFASQTDVTVAL 1188

Qy 1202 ----EENYATNHNIEGTSYALLALLMKKKFDDQGPVLRWLTQNFNGEYTGQATVMAF 1257

Db 1189 KALSEPALMNTERR-TNIOVTVTGSSPSPPL-----AVQOP--NAVNISANG 1232

Qy 1258 QALAEYEIQMPHKOINDITTELPRREVPIRYINENALLARVETFKLNODITVTASG 1317

QY 1233 PEPALCOLNVTYVNVKASGSSRRRSTIQNOEAPDLVAENKNDLNNVD-----LNNCT 1286  
 Db 1318 DKAATMTTLTFYNAQ-----QEKANVCN--PHLNVSE--NHLNAMAKGALMLKICT 1369  
 QY 1287 SFGSGRSGMALMEVNLISGFVNPSPALIS-----LSEYKAYEYDHGKLN-----LYLDS 1336  
 Db 1370 RVLGEVDSTMTLIDISMLTGLPDADLTRLSKGVDRYISRYEVDNNNAQKVAVIYIYLNK 1429  
 QY 1337 VNEIQ-FCVNIIPAVRNFVSTQDASVSIYDYEPBRQAVRSYNSKSLSCDLCSDVQG 1395  
 Db 1430 VHSSEDECHAFKILKIFEVGFIQPGSVKYSYNNDEKCTKRYHPDKGTGLLNKICIGNV 1489  
 QY 1396 CRPCEDGASGSHR-----HSSVIFIFCFKLLYFME 1425  
 Db 1490 CRGAGETCSLNQERIDVPLQIEKACETNNVDYVYKTLRIE 1533

RESULT 8  
 US-09-017-947-2  
 ; Sequence 2, Application US/09017947  
 ; Patent No. 6303754  
 ; GENERAL INFORMATION:  
 ; APPLICANT: VOGEL, CARL-WILHELM  
 ; APPLICANT: BREDEHORST, REINHORST  
 ; APPLICANT: KOCK, MICHAEL  
 ; APPLICANT: FRITZINGER, DAVID  
 ; TITLE OF INVENTION: RECOMBINANT PROCVF  
 ; NUMBER OF SEQUENCES: 39  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: OBION, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,  
 ; STREET: 1755 S. JEFFERSON DAVIS HIGHWAY  
 ; CITY: ARLINGTON  
 ; STATE: VA  
 ; COUNTRY: USA  
 ; ZIP: 22202  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: Patent in Release #1.0, Version #1.30  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/09/017,947  
 ; FILING DATE:  
 ; CLASSIFICATION:  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: US 08/662,227  
 ; FILING DATE: 14-JUN-1996  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: OBION, NORMAN F.  
 ; REGISTRATION NUMBER: 24,618  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: 703-413-3000  
 ; TELEFAX: 703-413-2220  
 ; INFORMATION FOR SEQ ID NO: 2:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 1642 amino acids  
 ; TYPE: amino acid  
 ; STRANDEDNESS: single  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: protein  
 ; US-09-017-947-2

Query Match 12.2%; Score 894; DB 4; Length 1642;  
 Best Local Similarity 22.5%; Pred. No. 4,4e-65;  
 Matches 361; Conservative 295; Mismatches 672; Indels 276; Gaps 56;

QY 18 ALAV-----APGRFLVAPGLIRPGAVTIGVELL-BHCRSOUTVAKA-----ELIKTAS 67  
 Db 10 ALLTLGPGSSHGALYTLITPAVLRTDTEQILVEAHGSDTPKQDIFVHDPKQKTLF 69

QY 68 NLTVSLAEVFEKSGFKTLTLPSLPLNSADEIYELRVGTODEILFSNSTRLSEFTK 127  
 Db 70 QTRVDMNAGMLVT---PTIEIPAKVESTSRQNVVGVTSQVLEFKVLLSYQSS 126  
 QY 128 RLSVFIQIDKALYKQKQVFRVITLPSDFEYKTSNLILIKDKPSNLIQOMLSQOSDLG 187  
 Db 127 --FLEIQTDKGIYTPGSPVLVRFVSMDNSTSRKNTVIVEOTPEGLIVS---SNSVTLN 181  
 QY 168 VISKTFOLSSHPLIGDMSI--OVQVNDQTYQSFOVSEYVLPKFEVTLQTP---LYCSMN 242  
 Db 162 FF-WPYNLPLDVLISGTMWIVAKIEHSPENTYATPDVRYKTVLPSFEVRLQPSKEFYIDGN 240  
 QY 243 SKHLNGITTAKYTGKPKYKGVDTLTFPLSPWKKKNI---TKTFKINGSANSPFNDEE 298  
 Db 241 --ENFHSITARYLAGEVEG--VAFVLFQVKIDDAKSIPLSLTRPIIDGOKATLKNDT 298  
 QY 299 MKNVMDSSNGISEYLDSSPGPVELLTVTSVSGISNNTVNFKKQHDYIIEFDYTT 358  
 Db 299 FRSRPPNINELVGHTLVAS-----VYVMTESGSDMVVTEQGIHYVASPYOHFTKTPK 352  
 QY 359 VLKSLNFTATVKTYPADGNQTLLEBRNNVITVTOGNTEYWSGNSGQKKEAVQKI 418  
 Db 353 YFKGMPELTVYVITNPDSPPA-----AHVNVSEAF-----HSMGTLSDGAKL 398  
 QY 419 NYTVPOSGTFKIEPPIEDSSELQKAYFLGSKSMVAHSLFKSPKTYIQLKTRDENIX 478  
 Db 399 ILNPLNMQ---SLPIVTRTHGDLPRERQATKSMWTAIAYTOGSGSGVYLVHVAITSTBIK 455  
 QY 479 VGS--PEFLVYSGN---KRLKELSYNVVSGQLVANGK-----QNSTMFSLPENSWT 526  
 Db 456 PGDNLPVNFVNVKGNANSLKQIKFTYLLINKKIFKYGQRPBGQNLVYNNMLHTTDPDI 515  
 QY 527 PKACIVYIYIDDEGISDVAKIPVQ-----LVFK--NKIKLWVKYKAPSEKVSIRI 578  
 Db 516 PSFRFVAYIQVGNNEIVADSVWVDKOTCKGTLLVVKGNLIQM-----PGAAMKTKL 567  
 QY 579 SVTOPDSIVGIIVADKSVNLNANASNDITMENNVHELBYNTGYLYGFMNSFAVFOECGL 638  
 Db 568 E-GDPAVGLVAVDKAVYVNDKRYKISQAKIMPTIEBDSGCTAGSQQNNLGVFEDAGL 626  
 QY 639 VWLTDANI--TYD----- 649  
 Db 627 ALTTSTNLTQKRSAAKCPQANRRSSVLLDSNASKAEPQDQDLKCCEDVMHENP 686  
 QY 650 -----YIDGVND--NAEYARFM--BENE--GHLYDIHDF 678  
 Db 667 MGYTCERAKYIQQGDACKAFLCCRYIKGVNDENORESELFLARDNDEGFLTADSDII 746  
 QY 679 SLGSSPHVKAPPETWIL-----DTNMGYRIYOEFEVTVPDSITSVAVTGFISEDIG 732  
 Db 747 S-----RSDPFKSMWLTDLTLEBRNSQISSKTMSFYLRDSITTVVLAASFPTPKG 799  
 QY 733 LGLTTTPEVLEAFQPFIFLNLPSYVIRGEEFALEITIFNYLKATERYKIIIEKSKEDI 792  
 Db 800 I-CVAEPEIRVMKVFIDLOMPYSVVNQEVEIRAILAHNVNBDIYVRVELLYNPAP-- 856  
 QY 793 LMTSSEINATGHOQTLVPSBDATVLEPIRPHLG--EITTYTALSPTASDAVTOML 850  
 Db 857 --CSASTGQRRYQFPKALSSRAVPVIVLEQGLHDEIKASVOBALMSDGRKKLK 914  
 QY 851 VKASGIEKYSQSLDLDTNRLQST-----LKTLSFSPPPNVTYSGSERVOITAIGVL 904  
 Db 915 VPEGVQVQSIVTYIKLDRAGVGVGTQLEVKARKLDRVDVDTIEFTKIIIIQSPVAQII 974  
 QY 905 GPSING--LASLIMPYGCGQNNMINPANI---YIIDLTKKKQLDNLNLEKALSEMRQ 959  
 Db 975 ENSIDGSKLNLILITPGCGQNNIRMAAPVIAITYVLDTEQMETLGINRTEAVNOIVT 1034  
 QY 960 GYORELLYQREDGFSAPAGNVDPGSGTMSAFVLRCP--LEADPIYIDQVNLHTTYTML- 1017  
 Db 1035 GYAQQMVYKADHSHYAAFTN--RASSSMWLTAYVVKVAPMAKMAVAGISHEIICGVRLI 1092  
 QY 1018 KGHQKSNGEFWDPGRVHISELOG--GNKSPVTLTAVTVSLGVR-----KYQPNID--V 1068



QY 851 VKAEGIEKYSOSILDLTDNLOST-----LKTISEFPPTVTGSRVOITAIQDVL 904  
 Db 915 VVEBQKSVITVTKLDPRAKGVGGQLEIVAKRKLDREVPTLETKIIOGDPAVQIT 974  
 QY 905 GPSING--LASLIRMPYGGGEONMNFAPNI---YLIDVLTKKQUTDLKFKALSPMQ 959  
 Db 975 ENSIDGSKLNHLITPSSGGEONMTRMAPVATYTLDTTEOMETIGINRRTAVANQIT 1034  
 QY 960 GYORELLYQREDSFSAFNGVNDPSGSTMVSAFVLRCE--LEADPIDIDQNVLRHTYTWL 1017  
 Db 1035 GYAQQMVYKKADHSYAAFTN--RASSWLTAVVYKFMAMAGVAGISHEIICGVRLI 1092  
 QY 1018 KKHQSNHGFMDPGRYHSELOG---GNKSPVTLTAVYTSILGR---KYQPNID--V 1068  
 Db 1093 INRQOPDGAFFKENAPVLSGTMOGIGAGAEFVYLTAFILVALLSEKTCINDYVNSIDSSI 1152  
 QY 1069 OESHFILSEFSEFSGISDNVTLTALITVYALSSVSPKAKELNMLTYRABEGEMQFVSS 1128  
 Db 1153 KKAITYLKKYER--IQRPYTLTAVALA-----AADLNDRVLMASGRDHV--- 1201  
 QY 1129 SKLSWOPRSLDIEVAAYALSHFLQFOTSEGIPIMRWLSRQNSLGGFASGTOPDTVAL 1188  
 Db 1202 ---BEVNAHTNIEGTYALALAKMKKFDOTGPIVWLTQNTYGETYQOTAWAF 1257  
 QY 1189 KALSEFALMNTER--TNIQVTYGPSSPSPL-----AVQP---NAVVISANG 1232  
 Db 1258 QALAEYEQIOMPHKDLNDITIELPDRVYIRYRINYEVALARVETELNDIYVTSAG 1317  
 QY 1233 GFALCOLNVVYVKAASSGSSRRRSIQNGEAPDLVAVKENDLNHD-----LVNCT 1286  
 Db 1318 DKAATMTITTFYNAOL---QEKAVCNK--FHLNVSE--NHLNANGAKAMLMKICT 1369  
 QY 1287 SFGSGRSGMALMEVNLISGFVWSEALS-----LSEYKAYEYDHGKLN-----LYLDS 1336  
 Db 1370 RYLGVDSTMTITIDISMLTGFPLDPAEDLRLSKVDRIYRKYEVUNNAQKAVIITLAK 1429  
 QY 1337 VNEIQ--FCVNIPAVENFKVSNTOASVSIIDYEPBRQAVRSYSEVKLSGDLCSDVQ 1395  
 Db 1430 VSHSEDECHFKILHFEVGIQPSVVKYSYNIDCKTKFYHDKGTGLINKICIGNV 1489  
 QY 1396 CRPCDGSAGSH-----HSSVFICFKLLYME 1425  
 Db 1490 CRACGETCSLHNERIDVPLQIERACETMNDVYVYKTLIRIE 1533

RESULT 10  
 US-08-662-227-35  
 ; Sequence 35, Application US/08662227  
 ; Patent No. 5922320  
 ; GENERAL INFORMATION:  
 ; APPLICANT: VOGEL, CARL-WILHELM  
 ; APPLICANT: BREDEHORST, REINHORST  
 ; APPLICANT: KOCK, MICHAEL  
 ; APPLICANT: FRITZINGER, DAVID  
 ; TITLE OF INVENTION: RECOMBINANT PROCVF  
 ; NUMBER OF SEQUENCES: 39  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,  
 ; ADDRESSEE: P.C.  
 ; STREET: 1755 S. JEFFERSON DAVIS HIGHWAY  
 ; CITY: ARLINGTON  
 ; STATE: VA  
 ; COUNTRY: USA  
 ; ZIP: 22202  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: Patentin Release #1.0, Version #1.30  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/662,227  
 ; FILING DATE: 14-JUN-1996  
 ; CLASSIFICATION: 530

ATTORNEY/AGENT INFORMATION:  
 NAME: OBLON, NORMAN F.  
 REGISTRATION NUMBER: 24,618  
 REFERENCE/DOCKET NUMBER: 1126-0107-0X  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 703-413-3000  
 TELEFAX: 703-413-2220  
 INFORMATION FOR SEQ ID NO: 35:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 1648 amino acids  
 TYPE: amino acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein  
 US-08-662-227-35

Query Match 12.0%; Score 880; DB 2; Length 1648;  
 Best Local Similarity 22.7%; Pred. No. 6.6e-64;  
 Matches 365; Conservative 295; Mismatches 664; Indels 284; Gaps 59;

QY 18 AALAV-----APGRFLVAPGIIIRPGSNVITGVELL--EHCPQVTVKA-----ELKLTAS 67  
 Db 10 AALLIGPSSSHGALYTLITPAVLAFTDTEQILVEAHGDSPPKQDLIFVHDFPRKQTLF 69  
 QY 68 NLTVSVLEAGVFEKGSFKTLTSLPLNSADEIYELVTRTODEILFSNSTRLSFETK 127  
 Db 70 QTRVDMNAGGLVY---PTIEIPAKVSTSRQNYVVGVTGQVRLKAVLSTQSS 126  
 QY 128 RISVFIQTDKALYKPKQEVKFRIVTLFSDPKRYKTSNLIL--KDPKSNLIQWLSSQSD 185  
 Db 127 --FLFIQTDKGIYTPGSPVLRYVFSM--DHHTSKNKTVIYVDFQPEGILVS---SNSVD 179  
 QY 186 LGVISTKQLSHPILTGMSI--QYVNDQYTYGSSFOYSEVYLKPEFTYLTQP---LYCS 240  
 Db 180 LNFV--WPNLNDLVSLGTRVAKYEHSPENVYAVFDRKYLVLPSFEVRLPQSEKPFYID 238  
 QY 241 MNSKHLNGTITAKTYVGRVADVTLTFLPLSFMGKKNI---TKYFKINGSANFSND 296  
 Db 239 GN--ENFHSITARYIYGEVEGS--VAFVLFYKIDAKSIDSILTRITIIDGKATILKR 296  
 QY 297 EEMKNVMSNGLSSEYLDLSSPGPVEILITVYESVTGISRVNSTVFPKQHDYIIEFDY 356  
 Db 297 DTFRSRFPNLNBLVCHTLVYAS-----VTWTEGSDMVTWTEGSDIHIVASPYQIHPKT 350  
 QY 357 TTVLAKPSLNTATYKVRADGNQLTLEBRNNVITVYQRYTERYSSGNSGNGQMEAVQ 416  
 Db 351 PKYFRPGMPYELTYVYNPDGSPA-----AHVPVSEAF-----HSMGTLISDGA 396  
 QY 417 KINYVPOSGTFKIEPILDESELOLKAYFLGSKSSMAVHSLFSPSPKTYIQKTRDEN 476  
 Db 397 KLILNIPLANQ---SLPTVATNHGDLPRERQATKSMTAIAIYOTGGSGNYLHAITSTE 453  
 QY 477 IKVGS--PELVVSGN-----KRLKELSYMVVSRQLVAVGKONS-----TM--FSLTPE 522  
 Db 454 IKPGDNLPVNNVAVGNANSLKQIKYFTYLLINKKIKFVAGQPRRDGNQVLTMLNHLITPD 513  
 QY 523 NSWTPKACVITYIYEDGEIISDVLKIPVQ-----LVFK--NKIKLWYSYKAPSESKV 574  
 Db 514 LIPSPR--FVAYYQGNNEIIVADSVWVDVKOTGCTLVVKKGDNILQMGAMKIKLJEDF 571  
 QY 575 SLRISVTPDSIVGVIVAVDSVNLNMASSNDITMENVVHLELNTGYLLGFMNMSFAYFQ 634  
 Db 572 GAR-----VGLVADKAVVYVLDNKYIKQAKIMDTLEKSDPGCTAGSGGQNNLVGE 622  
 QY 635 ECGMLVLTDAVL--TKD----- 649  
 Db 623 DAGALATSTYLNKQSSAAKCPQANRRRSSVLLDSNASKAAEPQDQLRKCCEDVM 682  
 QY 650 -----YIDGYVD--NAFYERFV--EENE--GHVVD 674  
 Db 683 HENPWGYTCERAKYIOGDACKAFLCCRYIYGVDEQNRSEBELFLARDNDNDEGFAD 742  
 QY 675 IHDPSLGSPIVRKHFPETWIL-----DTNMGYRTIYQEFVYVVPDSITSWAVTGFYIS 728

```

Db 743 SDIIS-----NSDFKMMMLTKDLTEEPNSGSSKSSKMSFYLRSDITTWVAVLAVSFT 795
Qy 729 EDLGLGTTTPVELQAFQPEFIPLNLPYVIRGEEFALETITFNLYLKADATEVKVILEKSD 788
Db 796 PTGDI-CVAEYELRWKVFPILOMWYSVKNQVEIRAILNHYVEDIYVVELLYNP 854
Qy 789 KFDILMTSSSEINATGHOQTLVSEDAATVLPFRPHLG--EIPITVALSPASDAVT 846
Db 855 AF-----CSASTKGORRQOPPIKALSSRAVPFVLPLEOGIHDEIKASVOEALMSDVR 910
Qy 847 OMLIVKAGIEKSSQSILDLTONRLQST-----LKTLSFSPPTVVGSEVQITAI 900
Db 911 KKLKVPBEGVQKSVITVKLDPPRAKVGQTOLEVIKARKDDRDVDTETIKTIIIOQDPV 970
Qy 901 GDVLGSPING--LASLIRMPYGCCEQNMINFAPNI---YLLDYLTKKKQOTDLKELAKS 955
Db 971 AQLIENSIDSKNLHLITPSGCEQNMIMAAVAVITVYIDTTECQETIGIRRTAVN 1030
Qy 956 FMRQYORELLYQREDGFSFAPGNYDPGSGTWSAFVLRCP--LEADPYIDIDONVLRHY 1014
Db 1031 QIVGVYQOQWYKKADHSYAAFTN--RASSMLTAAVYVKAAMAKKAVAGISHIICGV 1088
Qy 1015 TWL-KGHOKNGEWFMDGRVYHSELQ---GNKSPLYLTAYITVSLGVR---KYQPN 1066
Db 1089 RMLILNRQOPDGAKEKAPVLSGTGGIGQAESEVYLTAFITVALLSRTICNDYVNSL 1148
Qy 1067 D--VQBSIHFLSEFSRGISDNTYLTALITYALSVSQPKKEALMTWRAEOGQWQFV 1124
Db 1149 DSGIKRATNVLKRYEK-LQRPYTTALTAYALA-----AADQINDRVLAASITGSDHW 1201
Qy 1125 VSSESKLSDMOPRSLDIEVAAYALSHLQFQTSSEGIPIRMWLSRGRNSLGGFASQTDT 1184
Db 1202 -----EENYNAHNIIEGYSYALLLKKMKKFDQGPVIRWLTQNFYGETYQOAT 1253
Qy 1185 TVALKALSEPAALMNTER--TNIQVTVYGPSSPSPL-----AVQOP---NAVNI 1228
Db 1254 VMAFOALAEYELQPHKDLNLDITIELPREVPIRIRYENALARTVETKLMODITV 1313
Qy 1229 SANGFGFALCOLNVVYVKAAGSSRRRSIONOEAFDLDAVAKENKDLNHD-----L 1282
Db 1314 TASGDGAKTITLTFYNAQI-----QEKANVCNK--FILNVSVE--NHLNMAKAGKALML 1365
Qy 1283 NVCSTFSGPGRSGMALMEVNLGSGFWPSEALS-----LSETVKRYEDHGKLN-----L 1332
Db 1366 KICTRYLGEVDSWTITIDISMLTGFLPDADLTRLKSGVDRIYSRYEVDNMAQKAVI 1425
Qy 1333 YLDSVNETQ-FCVNTPAVRNFKVSNTOVASIYDYEPRQAVRSYNSVEYKLSGCDLCS 1391
Db 1426 YLKNVSHSEDECLPKLIKHFEGVFIQGSVKVSYNLDKCTKTFHPDKGTGLNKIC 1485
Qy 1392 DVQGRCEPDQASGSH-----HSSVIFFCFKLLYFWE 1425
Db 1486 IGVNCRCAGETCSLNHQERIDVPLQIEKACETNVDVYTKLRIEE 1533

```

```

RESULT 11
US-09-017-947-35
; Sequence 35, Application US/09017947
; Patent No. 6303754
; GENERAL INFORMATION:
; APPLICANT: VOGEL, CARL-WILHELM
; APPLICANT: BREDEHORST, REINHORST
; APPLICANT: KOCK, MICHAEL
; APPLICANT: FRITZINGER, DAVID
; TITLE OF INVENTION: RECOMBINANT PROCV
; NUMBER OF SEQUENCES: 39
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
; ADDRESSEE: P.C.
; STREET: 1755 S. JEFFERSON DAVIS HIGHWAY
; CITY: ARLINGTON
; STATE: VA

```

```

; COUNTRY: USA
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/017,947
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/662,227
; FILING DATE: 14-JUN-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: OBLON, NORMAN F.
; REGISTRATION NUMBER: 24,618
; REFERENCE/DOCKET NUMBER: 1126-0107-0X
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-413-3000
; TELEFAX: 703-413-2220
; INFORMATION FOR SEQ ID NO: 35:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1648 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-017-947-35

```

```

Query Match 12.0%; Score 880; DB 4; Length 1648;
Best Local Similarity 22.7%; Pred. No. 6,6e-64;
Matches 365; Conservative 295; Mismatches 664; Indels 284; Gaps 59;

```

```

Qy 18 AALAY-----APGRRLVAPAGIIRGCVNTIGVELL-EHCSQVYVKA-----ELKKTAS 67
Db 10 AALLIGPSSSHALTYLTTPALVLRDTBEQJLVEAHGSGTPEKQDIPVHDFPKKQKTLF 69
Qy 68 NLTVSVLEAGVEFEKSPKTLTLPSPLSADEIYELRATGRTODEILFNSNTRLSPEK 127
Db 70 QIRVDNRPAGMLVY---PTIEIPAKEVSTDSRQNYVVOYTGPOVRLKEKVLISYQSS 126
Qy 128 RISVFIQTDKALYKPKQEVKFRIVTLFSDFKPKYKTSNLILI--KDKSNLIQOMLSQOQSD 185
Db 127 --FLFIQTDKAGIYTPGSPVLYRVFSW--DHHSKMKKTYIVERQTEGILVS---SNSVD 179
Qy 186 LGVISTFQLSPIILIGMSI--QOVANDQTYQSGVSEYVLPKREYVLTQF---LYGS 240
Db 180 LNEF-WPYNLPDLVSLGTWRIVAKYHSPENYTAAYFVDRKYVLPSEFVRLQSEKEFFYID 238
Qy 241 MNSKHLNGTITAKYTYGKPEVKGDTLTPPLPSFMGKKNI-----TKTFKINSANPSFND 296
Db 239 GN-ENFHVASTIARYLYGEEVGG-VAFVLFGVKIDAKKSIPOSLRIPIIIDGKATLGR 296
Qy 297 EEMKNVMDSSNGLSLEYLSDSPPEVILITTVTESVTGISRNSTVNFPEKHDIYIEFPDY 356
Db 297 DTFRSRFPMLNELVGHITLYAS-----VTVMTESSGDVVTTEQSGIHIVASPOIHFTKT 350
Qy 357 TTVLKSPLNFATVYKTRADGNOLTEERRNNVITVYORANTEYVSGNSGNGQKEAVQ 416
Db 351 PKYFKGMEYELVYVYTNPDGSPA-----AHVPVSEAF-----HSMGTLLSDGTA 396
Qy 417 KINYVPOSGTFKIEPILEDSELOLKAYFLGSKSMVHSLFKSPKTYIOLKTRDEN 476
Db 397 KLILNIPLNAQ--SLPIYRTNHGDLPREQATKMTALAYOTOGSGNYLHVALTSTE 453
Qy 477 IKVGS--PEELVSGN-----KRLKELSYNVVARGQVAVGKQNS-----TW-FSLTPE 522
Db 454 IKRGDNLPVNPNVKNAGNSLKGQIKYFTYILNKGKIFKYGORBRDQOMLVYTNLHITPD 513
Qy 523 NSMTPRACIVYVYIEBDGIIISVLKIPVQ-----LVK--NKILYMSKYAAESEKY 574
Db 514 LIPSPR--FVAAYQVGNNEIVADSVWVDVDTCMGTLLVKGDNLIQMPGAMKIKLEGGDF 571

```



```

QY 575 SLRISTQSDSYGIYAVADKSVNLMASNDITMENVHELELYNTGYIIGMFNNSFAVQ 634
DB 572 GAR-----VGLVADVKAAYVINDKYISQAKIMDTIKSDFGCTAGSQNNIGVE 622
QY 635 ECGLWLTJLANL-TTDO----- 649
DB 623 DGLALITSTNLTNTKORSAKCPQANRRRRSVLLDSNASKAAEFQODLRKCEDVM 682
QY 650 -----YIDGYD--NAEYAEERFM--EENE-GHIYD 674
DB 683 HENPMGYTCERKAKYIQEGDACKAALIECCRYIKGRDENQRESEFLARDNEDGFILD 742
QY 675 IHDPSIGSSPHVRKHPETWIML-----DTMNGYRIYQEFVTVYDSTSWATGCVIS 728
DB 743 SDIIS-----RSDFPKWMLMTKDLTEEPNQGISSTMSFYLADSTITWVVLAVSPT 795
QY 729 EDGLGLTTTPVELAOFQPFITFLNLPYVIRGEAEALITTFNYLKATEVKVITKSD 788
DB 796 PTKGI-CVAEPYERIRMKVFFIDLOMPYSVVKNQVEIQAIAIHNYNVEDIYRVVELLNP 854
QY 789 KFDIMTSEINATGHOOTLVVSEEDGATVLPPIRPTHG--EIPITVALSPTASDAT 846
DB 855 AF-----CSASTGQRTKQOPPIKAUSSRAVPYIVLEQGLHDEIKASVQALMSDGV 910
QY 847 QMILVKAEGIEKYSQSILLDLTDNRLOST-----LKTLSFSPFNTYTSERVOITAI 900
DB 911 KKLKTVPEGVQSIYIVTKLDPRAKGVGTQLEVIKARLDLRVPDTEIETKIITIGDVP 970
QY 901 GVLGSPSING--LASLIRMPYCGEGEONMNFAPNI--YILDYLTKKKOLTNLKEKALS 955
DB 971 AOIIEISIDGSKLNLHILTPSGGSEONMIRMAFATATYLLDTTEEMETLGINRTEAVN 1030
QY 956 FRRQGYORELTYOREBDSFSAFGNDPQSGTMLSAPVLCF--LEADPYIDIDQNVHRY 1014
DB 1031 QIVTVGAQOVYTKADHSTYAFATN--RASSSWLTATVAVVFMAAMVAGISHEITICGV 1088
QY 1015 TWL-KGHOKSGNEFMDPGRIHSELOG--GNKSPYTLTAYIVTSLIGR----KYQPI 1066
DB 1089 RWLINRQOPDGAFKENAPVLSGTMGQIGABEEVYLFILVALLBESKTCINDYVNL 1148
QY 1067 D--VOESIHFLSEFSRGISDNYTALITVALSSVGSPPAKALNLTMRABOEGMOW 1124
DB 1149 DSISIKATNYLLKTYEK-LQRPYTTALTAYALA-----AADQINDRVLMASTGRDH 1201
QY 1125 VSESELSQSPRSLDIEVAAYALLSHLOFQTSRGITPMWLSQGRSLGFASTOT 1184
DB 1202 -----EYNAHTNIEGTSYALLALMKKFDQGTGPIVWMLTDQNFYGETYGTQAT 1253
QY 1185 TVALKALSEFALMNTER--TNIOVTYVGSPPPL-----AVVOP--MAVNI 1228
DB 1254 VWAFOALAYEIQMFPKDLNDITIELPREVPIYIRINENALLARVEYTKLANODIV 1313
QY 1229 SANGFQALCOLNVYVKAASSRRRRIQOEAFDLDAVAKENDLINVH-----L 1282
DB 1314 TASGQKATMTILTFYNAQL--QEKANVCNK--FHLNVSE--NHLNMGAKGALML 1365
QY 1283 NWCFSFSGRGSGMLMEVNLISGFVNPSEALS-----LSEYVKAYEYHGLN-----L 1332
DB 1366 KCTIRLGEVDSTMTIIDISMLTGLPDAEDLTRLSKGDYRISRREVENNNAOKAVYI 1425
QY 1333 YLDSVNETO--FCVNIPAVNFKVSNTOQASVIVDYBPRQAVRSYNSVNLSDLS 1391
DB 1426 YLANKSHSDECLHPIKLHGFVGFQPGSVKYISTYINLDEKCTKYHHDKGTGLINKC 1485
QY 1392 DVQGRPCEDGASGSH-----HSSVIFICFKLLYFME 1425
DB 1486 IGVNCRCAGETCSLNHGERIDVPLQIEKACETNVYVYKTKLIRIEE 1533

```

RESULT 12  
US-08-447-411-76  
Sequence 76, Application US/08447411

```

; Patent No. 5773243
; GENERAL INFORMATION:
; APPLICANT: FRITZINGER, DAVID C.
; APPLICANT: BREDEHORST, REINHARD
; APPLICANT: VOGEL, CARL-WILHELM
; TITLE OF INVENTION: DNA ENCODING COBRA C3, CVF1, AND CVF2
; NUMBER OF SEQUENCES: 81
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: P.C.
; ADDRESS: 1755 S. Jefferson Davis Highway, Suite 400
; CITY: Arlington
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/447,411
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/043,747
; FILING DATE: 07-APR-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: OBLON, NO. 5773243man F.
; REGISTRATION NUMBER: 24,618
; REFERENCE/DOCKET NUMBER: 1126-101-0
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 413-3000
; TELEFAX: (703) 413-2220
; TELEX: 248855 OPAT UR
; INFORMATION FOR SEQ ID NO: 76:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1333 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-447-411-76

Query Match 10.5%; Score 768.5; DB 1; Length 1333;
Best Local Similarity 24.4%; Pred. No. 9.7e-55;
Matches 301; Conservative 208; Mismatches 477; Indels 247; Gaps 45;

349 YIIFFDYTYVLPKPSLNFATVKTTRADGNQLTLEERNNVITVTOHNYTEWMSGNSG 408
DB 22 YQIYFTKPKPKFKPMPYELTYVYTKPDGSPAA-----HVPVSEAIHSE-----G 67
QY 409 NOKMAVOKINYVQSGFTKIEPFILEDSSSELQKAYFLGSKSMAYHSLFKSPSKTYI 468
DB 68 TLLSGTKLFLNTPQNAQ---SLPITVRYTHGDLPREROAIKSWTATAYOTQGGSGNYL 124
QY 469 OLKTDENIKVGS--PFLVLYSGN---KRLKELSYMVYVSGOLVAVGKONS-----T 515
DB 125 HVAITSTIKGDNLPVFNPNRGNANSLNQIKYFTYLLINKGKIFKVRORHGGDGNLVT 184
QY 516 M-FSLTPENSWTPKACVIVVYIEDDEGISDVLKIPVQ-----LVFK-----NKIRLY 562
DB 185 NMLHITPD--LIPSRFAIYQVGNNEIVADSVWVDVDTGCTGVYVAGATSRDRIQM-- 241
QY 563 WSKVAEPSEKYSRLISTQPDSTIGIYAVADKSVNLMASNDITMENVHELELYNTGY 622
DB 242 -----FGAAMKIKLE-GDPGAWIGLVAVDKAEVYINDKYISQAKIMDTIKSDFCT 293
QY 623 LGMFNNSFAVQEGSLWLTJLANL----- 647
DB 294 AGSGQNNIGVFEDAGLALITSTNLTNTKORSAKCPQANRRRRSVLLDSNASKAAEQ 353
QY 648 -----KDYIDGYDN-----AEYAEERFMENEG-----HIV----- 673

```

```

Db 354 DDLRKCCEDEGMENPMGHTCEKREKYLQEGDACKAAFLCECHYIKGIQDNKRESEFL 413
Qy 674 ---DIHFSLGSSPHV-RKHFPETIWL-----DTNMGYRIQEFVTVPDSITSW 720
Db 414 ARSDFEDDLFEGNITSRSDPFESWLMWOLSEHPNSKGISKIVPY---LDSITTW 470
Qy 721 VATGFVISEDGLGTTTT-----PVELQAPQPFIFLNLPSYVIRGEFALEITTFNY 773
Db 471 -----ELLAVGLSPTKICVAEPYEITVMKDFIDLOLPYSVYKNEQKIRAVLYNY 522
Qy 774 LKDATEVVIIEKSKPDIIMTSSEINATGHQOTLLVPSBDGATVLPFIRPHLG--EIP 831
Db 523 ADKDIYAVVELLYSPAFCSASTESQ---RYRQLPKALSSRAVSFVIVLEQGLHVE 578
Qy 832 ITVTALSPASDAVATQMLVKAEGIEKSYOSILLDLTDNLQST-----LKTLSFSEFP 885
Db 579 VTAASQGLMSDGVKKLKVPEGEWKSIVTIIELDPHTKIGIGTQVGLVANKLNDVRP 638
Qy 886 PNTVTSERVOITAGDVLGPSING--LASIIRMPYGGGEONMINF-APNT--YILDYLT 940
Db 639 DTEIETKITIQQDPVAQTENSIDGSKLNLIIFFPGGEQNMIRMTAPVATVYLDITQ 698
Qy 941 KKKQITNLKKALESFMRQIQRELLIQRDEGSFSAFENVPSGSTWLSAVLACF--LEA 999
Db 699 QMETLGIIRRTAVAVQIMTYAQQLVYKKAHSYAFTN--SASSWLTAVVYVIFALAA 756
Qy 1000 DPYIDIDQNVLRHTYTWL-KGHOKSNGEFPMDPGRYHSELQGNKSPV---TLRATVYTS 1055
Db 757 KIYVDINHEIYCGGKRWLLIIRQKTDGVRENAPVLFSTWQGGIQAEPBSGLTAFILVA 816
Qy 1056 LLGYRKTQPNIDVOESIHLESEFSG-----ISDNYTLALITYALSVSQSPKX 1105
Db 817 LLESRSI-----CNAYINILDSISISKANDVLLKYEKQLQRYTTALTAYALA-----AA 865
Qy 1106 EALNMLTRAEQEGMQFVWSESEKLSQWQPSRIDIVAAVYALISHELOQTSBGLPIM 1155
Db 866 ERLNDRVTLMAASTGRNMEBPNA-----HTNIEGTSYALALAKMKKPEAGHV 917
Qy 1166 RWLSRQNSLGGFASSTODTVALKALSFALNMTER-TNFOVTWGPSSPSP----- 1217
Db 918 QMLDQYGGYGGTQATVMMFQALAYEIQMFKHKLNDITIELPREVPRIRYRINY 977
Qy 1218 ---LAVQPMAYN-----ISANGFGALCOLNVVYVYASGSSRRRSRISQOEAFFDLVA 1269
Db 978 ENALLAQVETKLMEDFTVASAGQKATMTILTYNNAQL-----REDANVCNK--PHLDVS 1031
Qy 1270 V-----KENKQDLNHYDLNVCISFGSPGSGMALEVNILSGMPWSEALS--LSETVK 1321
Db 1032 VENYQNLNKEAKGAKGALKIKICTRYLGEVDSVTMTIIDVSMLTGFVPDTEDLTRLSKVD 1091
Qy 1322 K-----VEYDH-----GKLNLYLDSVNETQ--PCVNI PAVRNFKVNSTODASVINDYEP 1371
Db 1092 RYISMFETNNMAOKGVIYIYLDKVSHEDECLHFKILKHEVGFIOGSAKVVSYNLD 1151
Qy 1372 QAVRSYNSVEKLSCDLCSDOVQCPCEDEGAS 1404
Db 1152 EKCTKIYHPDARTGLNKICVGNVCRCAEFTCS 1164

```

```

RESULT 13
US-08-662-227-34
Sequence 34, Application US/0866227
GENERAL INFORMATION:
APPLICANT: VOGEL, CARL-WILHELM
APPLICANT: BREDEHORST, REINHORST
APPLICANT: KOCK, MICHAEL
APPLICANT: FRITZINGER, DAVID
TITLE OF INVENTION: RECOMBINANT PROCVF
NUMBER OF SEQUENCES: 39
CORRESPONDENCE ADDRESS:
ADDRESS: OBLON, SPIVAK, MCLELLAND, MAIER & NEUSTADT,
ADDRESS: P. C.

```

```

STREET: 1755 S. JEFFERSON DAVIS HIGHWAY
CITY: ARLINGTON
STATE: VA
COUNTRY: USA
ZIP: 22202
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/662,227
FILING DATE: 14-JUN-1996
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: OBLON, NORMAN F.
REGISTRATION NUMBER: 24,618
REFERENCE/DOCKET NUMBER: 1126-0107-0X
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-413-2220
TELEFAX: 703-413-3000
INFORMATION FOR SEQ ID NO: 34:
SEQUENCE CHARACTERISTICS:
LENGTH: 1333 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-662-227-34

Query Match 10.3%; Score 758.5; DB 2; Length 1333;
Best Local Similarity 24.4%; Pred. No. 6.6e-54;
Matches 301; Conservative 208; Mismatches 477; Indels 247; Gaps 46;

Qy 349 YIIEFPYTTVLKPSLNFITYVKTADGQNLLEERRNNVITYVQRTETWSSGNSG 408
Db 22 YQIYFTPTPKYFKGMYELTYVYTKDGPAA-----HVPVSSAHS-----G 67
Qy 409 NQKKEAVOKINTVYPOSGTFKIEFPILEDSESLQKAYFLGSSKSMVHSLFKSPSTYI 468
Db 68 TTLSDGAKLPLNTPONAQ---SLPITVRTNHDDLPREKAIISMRTAVTQGGSGNYL 124
Qy 469 QLKTRDENIKVGS--PEELVSGN---KRLKELSYVNSRGOLVAVGKONS-----TM 516
Db 125 HVALITSEIKGQNLPLNFVNRGANSLNQIKYFTYLILNKGIKFKVGRQHRGDLVTM 184
Qy 517 -FSLTPEKSTPACACVYIYIEDGEI-ISDVLKIPVQ-----LVYK-----NKIKLY 562
Db 185 NLHITPD--LIPSRFVAYVQVONNELEVADSVWVDKTCMGTLVVKGATSRDNRLOM- 241
Qy 563 WSKYKAEPSKVSRLISVTQPSIVGI VAYDKSVNLKMASNDITMEVVAHELEYNTGY 622
Db 242 -----PGAMKIKLE-GDPGAMIGLVAVDKAEVYLNDKYKISQAKIMWTIEKSDGCT 293
Qy 623 LGEFMNSFAVFOEGIMVLTDAFLT----- 647
Db 294 AGSGQNNLGFEDAGLALTTSTNLNTRKQBSAKCPQANRRRRSSVLLDSNASKAAQFQ 353
Qy 648 ---KQYIDGVYDN---AEYAEPRMEENEG-----HIV----- 673
Db 354 DDLRKCCEDEGMENPMGHTCEKREKYLQEGDACKAAFLCECHYIKGIQDNKRESEFL 413
Qy 674 ---DIHFSLGSSPHV-RKHFPETIWL-----DTNMGYRIQEFVTVPDSITSW 720
Db 414 ARSDFEDDLFEGNITSRSDPFESWLMWOLSEHPNSKGISKIVPY---LDSITTW 470
Qy 721 VATGFVISEDGLGTTTT-----PVELQAPQPFIFLNLPSYVIRGEFALEITTFNY 773
Db 471 -----ELLAVGLSPTKICVAEPYEITVMKDFIDLOLPYSVYKNEQKIRAVLYNY 522
Qy 774 LKDATEVVIIEKSKPDIIMTSSEINATGHQOTLLVPSBDGATVLPFIRPHLG--EIP 831
Db 523 ADKDIYAVVELLYSPAFCSASTESQ---RYRQLPKALSSRAVSFVIVLEQGLHVE 578

```





GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Comugen Ltd.

OM protein - protein search, using sw model

Run on: January 15, 2004, 18:13:05 ; Search time 44 Seconds  
(without alignments)  
6636.344 Million cell updates/sec

Title: US-10-020-095-4

Perfect score: 7348  
Sequence: 1 MGGPPLTAAHLCTCTAAL.....HSSVIFCFCKLLYFMEIWL 1428

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 762491 seqs, 204481190 residues

Total number of hits satisfying chosen parameters: 762491

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-Processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: Published Applications\_AA:  
2: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep:\*  
3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pep:\*  
4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep:\*  
5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pep:\*  
6: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep:\*  
7: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pep:\*  
8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep:\*  
9: /cgn2\_6/ptodata/2/pubpaa/US09\_PUBCOMB.pep:\*  
10: /cgn2\_6/ptodata/2/pubpaa/US09\_PUBCOMB.pep:\*  
11: /cgn2\_6/ptodata/2/pubpaa/US09\_PUBCOMB.pep:\*  
12: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB.pep:\*  
13: /cgn2\_6/ptodata/2/pubpaa/US10\_PUBCOMB.pep:\*  
14: /cgn2\_6/ptodata/2/pubpaa/US10\_PUBCOMB.pep:\*  
15: /cgn2\_6/ptodata/2/pubpaa/US10\_PUBCOMB.pep:\*  
16: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB.pep:\*  
17: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB.pep:\*  
18: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3360	45.7	665	US-10-108-260A-3396	Sequence 3396, Ap
2	1854.5	25.2	1508	US-10-369-493-5314	Sequence 5314, Ap
3	1850	25.2	1519	US-10-369-493-5313	Sequence 5313, Ap
4	1441	19.6	1474	US-10-292-081A-15	Sequence 15, Appl
5	1441	19.6	1474	US-10-052-817-2	Sequence 2, Appl
6	1440	19.6	1474	US-09-873-403-5	Sequence 5, Appl
7	1440	19.6	1474	US-10-292-081A-10	Sequence 10, Appl
8	1440	19.6	1474	US-10-292-081A-12	Sequence 12, Appl
9	1440	19.6	1474	US-10-292-081A-13	Sequence 13, Appl
10	1440	19.6	1474	US-10-331-496A-38	Sequence 38, Appl
11	1437.5	19.6	1508	US-09-756-247-4	Sequence 4, Appl
12	1433	19.5	1450	US-09-756-247-23	Sequence 23, Appl
13	1433	19.5	1451	US-09-756-247-25	Sequence 25, Appl
14	1428	19.4	1500	US-10-292-081A-9	Sequence 9, Appl
15	1421.5	19.3	1492	US-09-981-151A-10	Sequence 10, Appl

16	1419	19.3	1451	US-09-756-247-24	Sequence 24, Appl
17	1384	18.8	1500	US-10-316-253-267	Sequence 267, App
18	1384	18.8	1500	US-10-316-253-269	Sequence 269, App
19	1255	17.1	1285	US-09-925-301-1394	Sequence 1394, Ap
20	1157	15.7	1411	US-10-094-886-124	Sequence 124, App
21	1113.5	15.2	941	US-09-981-151A-39	Sequence 39, Appl
22	1109.5	15.1	1436	US-10-094-886-126	Sequence 126, App
23	1086	14.8	936	US-09-981-151A-35	Sequence 35, Appl
24	1080	14.7	898	US-09-981-151A-36	Sequence 36, Appl
25	1075	14.6	936	US-09-981-151A-37	Sequence 37, Appl
26	1059.5	14.4	931	US-09-981-151A-38	Sequence 38, Appl
27	960	13.1	1661	US-09-842-758-42	Sequence 42, Appl
28	956	13.0	1663	US-09-842-758-43	Sequence 43, Appl
29	934	12.7	1663	US-09-842-758-41	Sequence 41, Appl
30	931	12.7	1663	US-09-875-519A-22	Sequence 22, Appl
31	921	12.5	1602	US-09-778-927A-59	Sequence 59, Appl
32	894	12.2	1642	US-09-925-442-2	Sequence 2, Appl
33	880	12.0	1648	US-09-925-442-35	Sequence 35, Appl
34	804.5	10.9	751	US-09-981-151A-80	Sequence 80, Appl
35	796	10.8	1612	US-09-842-758-6	Sequence 6, Appl
36	758.5	10.3	1333	US-09-925-442-34	Sequence 34, Appl
37	748	10.2	1251	US-09-778-927A-58	Sequence 58, Appl
38	747	10.2	643	US-10-292-081A-11	Sequence 11, Appl
39	747	10.2	643	US-10-087-188-2	Sequence 2, Appl
40	738	10.0	912	US-09-756-247-29	Sequence 29, Appl
41	733.5	10.0	882	US-09-756-247-39	Sequence 39, Appl
42	719.5	9.8	528	US-09-764-876-20	Sequence 20, Appl
43	718.5	9.8	531	US-09-764-853-583	Sequence 583, App
44	689	9.4	134	US-10-160-162-139	Sequence 139, App
45	689	9.4	134	US-09-820-649-139	Sequence 139, App

ALIGNMENTS

RESULT 1  
US-10-108-260A-3396  
; Sequence 3396, Application US/10108260A  
; Publication No. US20040005560A1  
; GENERAL INFORMATION:  
; APPLICATION: HELIX RESEARCH INSTITUTE  
; TITLE OF INVENTION: NO. US20040005560A1 full length cDNA  
; FILE REFERENCE: H1-A0106  
; CURRENT APPLICATION NUMBER: US/10/108,260A  
; CURRENT FILING DATE: 2002-03-27  
; NUMBER OF SEQ ID NOS: 5458  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 3396  
; LENGTH: 665  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-108-260A-3396

Query Match 45.7%; Score 3360; DB 12; Length 665;  
Best Local Similarity 100.0%; Pred. No. 4.9e-268;  
Matches 657; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	MGGPPLTAAHLCTCTAALAVAPGRLVTAAGIIRFGQNTTIVELLEHCPSQVTVYA	60
DB	1	MGGPPLTAAHLCTCTAALAVAPGRLVTAAGIIRFGQNTTIVELLEHCPSQVTVYA	60
QY	61	ELTKASNTVTVVAAEGVFEKGSFKTLTLPPLNSADEIYELVGTGDEILFNSNT	120
DB	61	ELTKASNTVTVVAAEGVFEKGSFKTLTLPPLNSADEIYELVGTGDEILFNSNT	120
QY	121	RLSFTKTSIVFIQTKALYKPKQEVKFRIVTLFSDFPYKTSNLTILKDPNSNLTQWL	180
DB	121	RLSFTKTSIVFIQTKALYKPKQEVKFRIVTLFSDFPYKTSNLTILKDPNSNLTQWL	180
QY	181	SOOSDLGVISKTFOULSHPIIGDWSIQVANDQTYGQFVSEYVLPFEVTLQTPLYCS	240
DB	181	SOOSDLGVISKTFOULSHPIIGDWSIQVANDQTYGQFVSEYVLPFEVTLQTPLYCS	240

```

Qy 241 MNSHNLNGTITAKTYGKPVKGDVTLTFLPLSFNGKKNTTKTEKINGSANFSNDEEMK 300
Db 241 MNSHNLNGTITAKTYGKPVKGDVTLTFLPLSFNGKKNTTKTEKINGSANFSNDEEMK 300
Qy 301 NMVSSNGLSEYLDLSSPGVEILLTTVESVTGISRANSTVFPKHQHYIIIEFDYTVL 360
Db 301 NMVSSNGLSEYLDLSSPGVEILLTTVESVTGISRANSTVFPKHQHYIIIEFDYTVL 360
Qy 361 KPSINFTATVATADGNQLTLBERNNVITVQRYNTEYWSGNSNGNQMEAVOKINY 420
Db 361 KPSINFTATVATADGNQLTLBERNNVITVQRYNTEYWSGNSNGNQMEAVOKINY 420
Qy 421 TVPOSGTFKIEPPIEDSSELQKAYFLGSKSMAVHSLFKSPKTYIQLTRDENIKVG 480
Db 421 TVPOSGTFKIEPPIEDSSELQKAYFLGSKSMAVHSLFKSPKTYIQLTRDENIKVG 480
Qy 481 SPFELVSGNRLKELSTVWVSRQGLVAVGKNSMTFSLTPENSWTPKACVIYIIEBDG 540
Db 481 SPFELVSGNRLKELSTVWVSRQGLVAVGKNSMTFSLTPENSWTPKACVIYIIEBDG 540
Qy 541 EISDVLKIPVQVFNKIKLYMSKVKAPESEKYSLSRISTOPDSIVGIVAVDKSVNLMN 600
Db 541 EISDVLKIPVQVFNKIKLYMSKVKAPESEKYSLSRISTOPDSIVGIVAVDKSVNLMN 600
Qy 601 ASNDITMENVVEHELELYNTGYLLGFMFMSFAVFOEGGLMVLTDANLTQDYIDGYDN 657
Db 601 ASNDITMENVVEHELELYNTGYLLGFMFMSFAVFOEGGLMVLTDANLTQDYIDGYDN 657

```

# RESULT 2

```

US-10-369-493-5314
; Sequence 5314, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; PRIOR FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 5314
; LENGTH: 1508
; TYPE: PRT
; ORGANISM: Caenorhabditis elegans
US-10-369-493-5314

```

```

Query Match 25.2%; Score 1854.5; DB 12; Length 1508;
Best Local Similarity 31.8%; Pred. No. 4.2e-143;
Matches 476; Conservative 292; Mismatches 562; Indels 169; Gaps 44;
Qy 15 VCTAALAAAPGRPLVTAAGIIRPGANTTIGVELLEHSPQVYKAEILKTASMLTYSVL 74
Db 28 VSTTAAPKPA-TYMLVAPAVRRPDQPSVCMNLKQATDEDMVRLEVRTERETIAAR 86
Qy 75 BABGFEKSGFKTTLPLPLNSAD--EIEELATVGRTO--EILFNSSTRLSFETKISV 131
Db 87 VISNL-KGIACTYSLSEMPAOSLTPROSTYKLYRGETLNABELFENENELKYQKALSV 145
Qy 132 FIQTDKALYKPKQEVKFRIVTLFSDFKPKYKSLNLIKDKPSNLIQOMLSQOQSLGYISK 191
Db 146 FIQTDRAIRPASIYRAIVKSDLKPYGNATIKIPDPSRNLIQITIGVTLDRGVYSG 205
Qy 192 TFOUSSHHIIGDMSIQV---VNDQYYOSFOVSEYVLPKFEVTLQRPYCSMNSKHN 247
Db 206 ELQLAETLLGDWLEFVETNSGVODKS---SFTVDYTVLPKFEVNIKTSFITTND-DLS 261

```

```

Qy 248 GTTAKTYGKPVKGDVTLTFLPLSFN-----GKKON---ITKTKIN--GS 289
Db 262 VFVDAKTYGKGVAGKAKYSLLELPMHRMAVPTIIDENGVKKEEBELMVERTKYLNQGE 321
Qy 290 ANSFNDEEMKNMDSNGISEYLDLSSPGVEILLTTVESVTGISRANSTVFPKHQHY 349
Db 322 AAIVFSNDELKR-----HKLHEWGGSIIIVASVEDITEIENNAHQISTREEV 373
Qy 350 IIEFDYTVTLKPSLNETATVATVTRADGN--QTLBERNNVITVQRYNTEYWSGNS 407
Db 374 KLDVEKGDFFKGLTYNNVVALKQMDPTVKATLPR-----VQSFYFVYP--YNNHDS 427
Qy 408 GNQMEAVQKINTVPOSGTFKIEPPIEDSSELQKAYFLGSKSMAVHSLFK----- 461
Db 428 SLQSEKTKIIVEVDAGHTSVLTQPPINCTSARIEAH-YDIGKDNFATPATSILVEA 486
Qy 462 --SPSKTYIQLTRDEN-IVGSPFELVSGNRLKELSTVWVSRQGLV-----ANGKON 513
Db 487 AVSEPTSLQLLDNEADVGVKSLSFSLKATQPLSTIYQVMSRSNIVVSQOMTVNSEH 546
Qy 514 STMSLTPENSWTPKACVIYV-YIEDGELISDVLKIPVQVFNKIKLYMSKVKAPESE 572
Db 547 ATI-SPPATNMAPKSLIYALIESQEVVLADPKVEGIRQNGVALSIDQAVEPQG 605
Qy 573 KSLRISVTOPDSIVGIVAVDKSVNLMNMSNDITMENVVEHELELYNTGYLLGF----- 626
Db 606 NVAFKVT-SPKNSFVGLVVDQSVLLKTGNDITREKVEQDLBNYDSNNVGGFGGPRPW 664
Qy 627 -----MNSFAVFOEGGLMVLTDANLTQD-----YIDGYDNA 658
Db 665 EALDRKRSIMRPMWIGSGSDAOSIFSNAGLVLTALLVREPQRFMSVMMMDGARGMA 724
Qy 725 EAA--FAAPMG-----GSSPPPTVYKFFPHTWMSLN-STSEVEMEIEAPD 771
Qy 716 SITSWATGVVISDGLGLTTPVBLQAFQPFPIPLNLYSVIRGEELAITTINYLK 775
Db 772 TITSWASTAINEENGGLGVAPTTSKLRVRRPFIQLNLPYAVRGEKFPALLVLFVYME 831
Qy 776 DAEVAKIT--EKSQKEDILMTSET---NATHQOTLLVPSDGN---VLEPIRPTH 827
Db 832 KEQDVTYLLKXDQSGDLKKGDTVVRDEVOQNVRIYSVAGGSGKAVPEPIYESSI 891
Qy 828 GEIPITVLTALPTASPAVNTQMLYKAELEKYSQSILDLTDNRLOSTTKTSFSPPN 887
Db 892 GEIPVHISALASQGDVAENMLAVDPQGYVDNIPFVIDLNNSDFG-KNELIWPND 950
Qy 888 TVTGSERVQITAGDVLGSPSINGLASLIRMPYCGGQNMINFAPNITLYDLTRKKQLTD 947
Db 951 VVDGSGKARLDVIGDMWGPTLVNNAHKLQMPYCGGQNMNLVPLNLYCYLRATNRNES 1010
Qy 948 NIKKALSPKQGYQRELYQREDGSPSANGVNDPSGTYLAFVLRCLFEADPPYDIDQ 1007
Db 1011 QLETKALKFLEQIQRELYTKRADNSFASAGDSKSTWLTTFVARSFHAHQYAFVDP 1070
Qy 1008 NVLHRTYTLKGQKSGSEGFMDPGRVYHSELQGG-KNSPVTLYRAYIVTSLGKRYQPN 1066
Db 1071 NVISRAVAFINSQOMESGAPAEERGEVHHKDMQGAODGVALLAFYLISIL-----ENGK 1125
Qy 1067 DVQESIHFLSEFSRGSIDNYLTALITYALSIVGSPKAKALMMLTWR--AEQEGMQFW 1124
Db 1126 ENKRAVLYLEKHLDEVSGNATYMAVAVALQLKSKQAGAFENLKGKIKVERSGDVKFA 1185
Qy 1125 VSSE--SKSDS-----WQPSLDIEVAVALLSHFLQFQTSSEGIPIYRWLSRQNSLIGF 1178
Db 1186 SAQKQVEKLEESRAYMFOAPVDIETTSYAVLSYLQNMQTSSELSIIRMLVVSQRNBLGCF 1245
Qy 1179 ASTQDTTVALKALSEFALMTERTNIOYTV--GSPSPS-----PLAVYQPM----- 1224
Db 1246 TSTQDTTVALQALSSTAAYVYSKHTSQVITLNGKHTSHDINRAIVLQSQSLND 1305
Qy 1225 AVNISANGFPAICQINVVYVVAAGSSRRRRRSIQNQEAFDLQVAVENKODLNHVDLVN 1284

```



Db 1306 AVSINAGGVFAQJLSYSTS-----YRSLNDAPFCSCQETKEIRAG-NRQLDLD 1355  
QY 1285 CTSFSGRSGMALMEVNLISGFMPVSEALISLSEYK-----KVEYDHGKLNLYLDSVNE 1339  
Db 1356 CCNTRPFGSGNMALAEIDALSGRFADEQVHTLTSIEDLQRYEMEDKDTKMYENFPLGG 1415  
QY 1340 TQPCVNI PAVRNFKVSNTODASVSIYDYEPRRAAVRSYNSSEVKLSCDLCSDVQGRP 1398  
Db 1416 RPYCJLSYSDVTYQVADQKRPANFRLVDYDPEEQLMKTYAAKQTRSLQKCG--EDCWP 1472  
RESULT 3  
US-10-369-493-5313  
Sequence 5313, Application US/10369493  
Publication No. US20030233675A1  
GENERAL INFORMATION:  
APPLICANT: Cao, Yongwei  
APPLICANT: Hinkle, Gregory J.  
APPLICANT: Slater, Steven C.  
APPLICANT: Goldman, Barry S.  
APPLICANT: Chen, Xianfeng  
TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF  
FILE REFERENCE: 38-10(52052)B  
CURRENT APPLICATION NUMBER: US/10/369,493  
PRIOR FILING DATE: 2003-02-28  
PRIOR APPLICATION NUMBER: US 60/360,039  
NUMBER OF SEQ ID NOS: 47374  
SEQ ID NO 5313  
LENGTH: 1519  
TYPE: PR1  
ORGANISM: Caenorhabditis elegans  
US-10-369-493-5313  
Query Match 25.2%; Score 1850; DB 12; Length 1519;  
Best Local Similarity 31.6%; Pred. No. 9.9e-143;  
Matches 477; Conservative 290; Mismatches 563; Indels 180; Gaps 44;  
QY 15 VCTAALAAVAPGPFVLTAGCIIRPGSNVITGVLELHCPQSQVTVKAKLLKTNASNLTVSVL 74  
Db 28 VSTTAPAVKPA-TYMLVAVAVRPPDPFSCVNLMLQATDEDMIVRIEVRTERNETIAR 86  
QY 75 EAGVEFGSKTLTLPSP.LPNSAD--EYELRVGRTOD-ETLFSNSTRLSSETRISV 131  
Db 87 VTSNL-KPGIAQVLSSEMPAOSLTPROSKLYTRGETNAELIFENENELKYDQALSV 145  
QY 132 FIOTDRAKYKQKQEVFRITVTLFSDFPKYKTSINILIKDKPSNLIQOMLSQOSDLGVISK 191  
Db 146 FIOTDRAIYRPAVLAVRAIVKSDLKFPYGNATIKIFDPSNRLISQITIGVTLDRGVSG 205  
QY 192 TFOGSHPLIGDMSIVQ---VNDQTYOSFOVSEYVLPRKEVTLQTPLYCSMNSKLN 247  
Db 206 ELQLAETTLGDMFIEVETSNVQVDS---SFTVDYTVLPKFEVNIKTSSFTIIND-DLS 261  
QY 248 GTTLATYTGKPKPGVTLTF--LELSFW-----GKKKN-----ITKTFKLN--GS 289  
Db 262 VFIDAKYTTGKGVAGAKVSLPWRHMAVPTIIDENGVKKEELMVERTVKLRQGE 321  
QY 290 ANFSFNDSEMKVNDWSNGLSEYLDSSPGVAILTTVTESVIGISRNVTNVFFKQHDY 349  
Db 322 AAVVFENDELK-----HKLHMGSGSIRIVASVTEDEITERNNATHQISTREAV 373  
QY 350 IIEFPDYTVLKPSLNPLATVKTADGN--QITLEERNNVITVYQNTYETYSNGNS 407  
Db 374 KLDVEKQGTJFKGLTYNNVVALKQMDPVPVKATLEPKR---YQVSTFYNP--YNHDS 427  
QY 408 GNGKMAVOKINTVQSGTFKIEFPILEDSSSELQKAFYLSKSSMAHSLFK-----461  
Db 428 SLOBEKETKIVEVDAGTSLVTLQPIINCTSARIEAH-IDIGGKDKFTKPTPIYSSLYEA 486  
QY 462 --SPSKTYIQLKTRDEN-IKVGSPELIVVSGNRKLELSYMWVSRQLV-----AVGKON 513

Db 487 AVSPKSKFLQLLADNEGADVQKSLSPSLKATQPLSTITTYQVMSRNVISQOMTVNSEH 546  
QY 514 SMPFLTEPNSMTWTPACIYV-YIEDDEIISDVLIKIPVQVLPKKIKLYMSKYAESE 572  
Db 547 KTT-SFPATNAPKSRILYVAILSSQEVLDALDFVEBGLPQVQVSLDKQAVEBQ 605  
QY 573 KVSLSISVQPSIYGVAVDKSVNLNMAASNDITMENVHELELYNTGYLGMF-----626  
Db 606 NKPKFVLT--SDKNSFVGLLVNQSVALTLKNGDITREKYEQDLENDSNNVGGFCGPRPW 664  
QY 627 -----MNSPAVQEGGLMVLTD-----NLT 647  
Db 665 EADIRKESIRPMMWIGGSDAQSIIFSNAGLVLTLDLALYREPOREFMSERRLTPGLLT 724  
QY 648 KDYIDGVYDNEAYARFEBNEBGHVLDHDSLSGS---PHVRKHFPETWITLDTNMGYR 704  
Db 725 VMMMDGAPGMAEA--FAAPMG-----GSSPPPTVRKFFPHTWIMSDLN--STS 771  
QY 705 IYQEPVTPDPSITSMVATGFVISEDGLGLTTPVELQAFQPFIFPLNLYSVYRGEEF 764  
Db 772 GEVMEIAPDITTSVASTPAINBENGGLVAPPTSKLRVRRPFIQNLTYAARRGKF 831  
QY 765 ALKITIYVLAQAEVKVIL--EKSDKFDILMTSSEI--NATGHQOTLVPSEDAAT--817  
Db 832 ALLVAVFVYMEKEQDVTLTKYDKDSGYDLKKDGTVVRDEVGQVNRIVSVAGGTSK 891  
QY 818 -VLEPIRTHGEPITVITALSPTASDAVQTMILYKAEIGKYSQSITLLDTNRLQST 876  
Db 892 AVEFPVSSIGELPVAHSALASQGGDAVENKLRLDPGQYVDRNIPVILDNNSDPS 951  
QY 877 LKTLSPSPPTVTVGSEBVOITAIQDVLGPGSINGLASLIRMPYGGGEONMIFANVITL 936  
Db 952 -KNLELIPNDVVDGSCARLDVIGDMGVPVLANNAHLVQMPYGGGEONMLNLPNITLV 1010  
QY 937 DYLTKKQITDNLKREKALSFMKQGYRELILYQREDSFSAFGANDPSSGTYLSAFVLACF 996  
Db 1011 KILRATNRESQLETKAIKFTIEQIGRELTYKRADNSFSAGDSDKASTWLTAVVASF 1070  
QY 997 LEADYITIDQNVLTARTYTLKGHOKSGEEMWDERVTHSELQGG--KNSPTLTLYITS 1055  
Db 1071 HHAKQYAVDPVDPVSRVAFVNSQMGSAFAEREVHHKMDQGGAGVALTAFVLIS 1130  
QY 1056 LLGYRKYQPNIDVQSHIFLSEPSRGSIDNVYTLALITYALSSVSPKAKELNMLTR- 1114  
Db 1131 IL-----ENGEMGKAVITYLEKHLDVSGNAYTMVAAYALQAKSKQAGAFENLKXK 1185  
QY 1115 -AEQEGNQFVWSE--SKLSDS---WQPSRLIEVAAYALLSHILOFQTSBGPIMRW 1167  
Db 1186 IYKSGDVYKFAQAQKVEKLKESRAYMQRAPVIDETTSYAVLSYLAQNGSSELSIRW 1245  
QY 1168 LSRQNSLGGFASITDITVALKALSEFALNTERNTIQVTV--TPSSPS-----PLAV 1220  
Db 1246 LVSCNNELEGFSITDITMAQALSSYAALVYSDKHTSQVITLNGKHTSHDINRMAIV 1305  
QY 1221 VQPM-----AVNISANGFPAICQALNVVNVVAVASGSSRRRRSIOQNEADLDVAVEN 1273  
Db 1306 LQSYLSSLDNVAISINANGTVVFAQLSYST-----YRSLNDAPFCSCQETKEI 1356  
QY 1274 KDLNLHVLDINACTSPGSGMALMEVNLISGFMPVSEALISLSEYK-----KVEYDHG 1328  
Db 1357 RAG-NRQLDLCNCTRKSGNMALAEIDALSGRFADEQVHTLTSIEDLQRYEMEDKDT 1415  
QY 1329 KLNLYLDSVNETQFCVNI PAVRNFKVSNTODASVSIYDYEPRRAAVRSYNSSEVKLSCD 1388  
Db 1416 KKNVVFENDELK-----HKLHMGSGSIRIVASVTEDEITERNNATHQISTREAV 1475  
QY 1389 LQSDVQGRP 1398  
Db 1476 KCG--EDCWP 1483  
RESULT 4

US-10-292-081A-15  
; Sequence 15, Application US/10292081A  
; Publication No. US20030162202A1  
; GENERAL INFORMATION:  
; APPLICANT: Kenneth David Becker  
; APPLICANT: Genul Velicelebi  
; APPLICANT: Xin Wang  
; APPLICANT: Randolph E. Tanzi  
; APPLICANT: Lars Bertram  
; APPLICANT: Aleister J. Saunders  
; TITLE OF INVENTION: SINGLE NUCLEOTIDE POLYMORPHISMS AND MUTATIONS ON ALPHA-2-MACROGLOB  
; FILE REFERENCE: 37481-3323  
; CURRENT APPLICATION NUMBER: US/10/292, 081A  
; PRIOR FILING DATE: 2002-11-08  
; PRIOR APPLICATION NUMBER: 60/337434  
; NUMBER OF SEQ ID NOS: 15  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 15  
; LENGTH: 1474  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-292-081A-15  
Query Match 19.6%; Score 1441; DB 12; Length 1474;  
Best Local Similarity 29.2%; Pred. No. 5.5e-109;  
Matches 453; Conservative 264; Mismatches 551; Indels 284; Gaps 60;  
12 LCCCTAALAVAPGRFVLTAPGIIIRPGNAVITGVLELHCPQVYVYKAEILKASNLTV 71  
15 LVLVLPDASVSGKQYVAVPSLHT-ETTEKGVLLSYLNETVYVYASLESVGRKSL 73  
72 -SVLAEG-VEKSGFCTLPLSLPLNGADE---IYELRVYGRTODELFSNSTRLSPET 126  
74 FTDLAENDVLHCAAF-----AVPKSSNEVWFVLYVQVGPQOE---FKKRTVWAKN 124  
127 KRISVFIQDALKYKPKQEVKRIYVLTFSDFKPYKTSNLIL-IRDKSNLIQWLSQSD 185  
125 EDLIVFQDTSIKYKPGQTVKFRVYVMDENFHLNLELPLVYIDPKGNRIAQWQSFQLE 184  
186 LGVISTKQLSHPLSLGMSIQVQVND--QTYVQSFQVSEVVLKFEVTLQTPYLCSSNS 243  
185 GGLKQFSPSLSEPPQSGYKVVQKSGGRTGH-PFYVEFVLPLFEVQVTPPKITILE 243  
244 KHLNGTIPAKTYGKPVGDTVTLPLPL-----SFMGK 276  
244 EEMVSVGGLYTYGKPVGHVTVSICRKYSDASDCHGDSQAFCEKBSGLNSHCFTYQ 303  
277 KKNITKTFKINGSANFSFNDEMNQVMSNGLSEYLDLSSPGVEILT-----TVTESV 331  
304 VK--TKVFLQKR-----KEYEMK-----LHTBAQIQEGTVVELTGRQSEIRTRI 347  
332 TGISRNVSTNVFFKQHDYIIIEFFDYTVLKPSLNFATVKTTRADGNOLTLSERRNVYI 391  
348 TKLS-FVVDVSHFRG--IPEFG-----QVRLVDGKGVPIP---NKVI 384  
392 TVTQRNTEVYSGNSGNQKAEAVQ-KINYT--VPOSGTFKIEPP-----ILSDS 438  
385 FI-RGNENANY--SNATTDHGLVQFSINTNNMGTSLTVRVANKDRSPCYGVWSEBH 441  
439 SELQKAYFLGSKSMVAHSLFKSPSKTYIQLKTRDENIKVS-----PFLVVSNGK--R 492  
442 EEAHTYV-----LVFSPSKSFVHLERPSHLEPCGHQVQYQAHYIILNGTLLG 489  
493 LKELSL--YMVVSRQQLVAVG-----KQNST---MESLTPENSMTPKACVIVYIEDDG 540  
490 LKKSIFYLLIMAKGIVRTGTHGLLVKQEDMKGHFSISIPKXSDIAPARLLIYAVLPTG 549  
541 EIIIDVLKIPVQVLPKKNIKLYMSKVKAPESEKXSLKISVQPSIYIGIYAVDSVNLN- 599  
550 DVIDSAKYDVENCILANKVDLSFSPSOLPASHHLRLRYTA--POSVCLARVVDSDVLLMK 608  
600 -----NNS-----NDITMENVVHLELYNTGY--YLGMFMNS-----FA 631

Db 609 PDAELSSASSVYNNLPEKDLTGPPGLNDQDNEDCINRHNHYINGITYTTPVSSITNEKDMYS 668  
Qy 632 VPOEGCLMVLTDANLTQDYIDGYDNAE-----YAEFMEENEGHIYDIHDFSL 680  
Db 669 FLEDMGLKAFNTSKIRKPKKCPQLQOYEMHGPBGKLVGFYESVWGRGARLVHVE-- 725  
Qy 681 GSSPH--VKHFPETWIMDTMGYRIYQEFVTVPDSITTSVWATGVVISEDGLGLTT 737  
Db 726 ---PHETVAKYFPEETWIMDLVVNSAGVAVGVTPDITTEKKAFCISEDAGGIS 782  
Qy 738 TPVELQAPFPFLNLPVSVINGEPALEITTFNLIKATREKXVILKSKDKDILMTSS 797  
Db 783 T-ASLRAFQFPFVELTMPYSVIRGEAFTLKATYANTLPCKIRSVOLDEASPAFLAVER 841  
Qy 798 E-----INATGHQOTLLVPEDEGATVLPFRPHLEIPIYVYALS-----PTA 841  
Db 842 EQAPHCICANGRO-----TVSMVTPSLGNVNTVABALESOLCTGEVSV 890  
Qy 842 ---SDAVTQMLVVAEGIEKSYQSILDLTDNRLQSTLTKTSFSPFPNTVGSERVQ 896  
Db 891 PEHGRKOTVAKPLVPEGELEKETTNSLCPGSGEVS---ELSLKLPNVVVEESARAS 947  
Qy 897 ITAIGDVLGPSINGLASLIMPYGCEBQNMNINAPNIYIIDIYTKKQLTDLNLERALSF 956  
Db 948 VSVLGDILGSAMQNTNLLQMPYGCGBQWVLPAPNIVYVDLYNETQQLTPREIKSKAIGY 1007  
Qy 957 MRQGYRELLYQREDSFSAFC-NYDPS--GSTWLSAFVRCFLEADPYIDIDQNVLHRTY 1014  
Db 1008 LNTGYQQLNPKYIDQSYSTFGERYGRNQNTVTLVFLVYQALAYFIFDAHITQAL 1067  
Qy 1015 TWLKGHQSNGEFPWDPGRVHSELQGNKSPVLTAVYVTSLLGYRKYQPNIDVOESIH 1074  
Db 1068 IWLISQKQKQKCFRRSSGSLNNMAIKGVEDEVLTSLAITALLEIPLTVTHPVYRNALFC 1127  
Qy 1075 LES-----EFSRGISDNTYLLIYALSVS-PAKAEALNLTWRAEOGGMQV--- 1124  
Db 1128 LESAMKTAQBGDGH-SHYVYKALLAFAFALAGQDKEVLSLNEBAVKKDMSVHERP 1186  
Qy 1125 VSESKLSDSMOPR--SLDIEVAVALLSHFLQFO---TSEGI-P---IKWLSRQNSL 1175  
Db 1187 QKRPAPVGHFHEQDAPAEEMTSYVLLA-YLTAQPAPISEDLTSATNIVKMTIKQONAO 1245  
Qy 1176 GGFASQDTPVAKALASEFALNTERNTIQTVTGSPSPPLAVVO----- 1222  
Db 1246 GGFSSQDTPVVALHSLSKYGAFTT-RTGKAQVTTIGSSGTFSSKQVQVNNRLLIQVS 1304  
Qy 1223 ---PMAVNISANGFPAICQLNVYTNVAKSGSSRRRRSIQONDAFDLVAVK---ENKD 1275  
Db 1305 LPELPGESYMKVYBEGCVYLDTSLKATYNI-----LPEKEFPFALGVQTLPTQCD 1353  
Qy 1276 DLN-HYDLANCTFSFGP---RSGMALMEVNLISGMVPEALST---SFTVKVVEYDH 1328  
Db 1354 EPRKHTSFQISLSVSTGSSASNMALIVDKWVSGTIPLKPTVKMLERSHVSRTVESSN 1413  
Qy 1329 KNLVYDSVNETQFCVNIIPAVRNFKVSNQDASVSIYDYVEPRQAVRSYS 1380  
Db 1414 HVLITDKNSNQLSLPFTVLDQVPRDLKRALVKKYDYIETDEFALAEYNA 1465  
RESULT 5  
US-10-052-817-2  
; Sequence 2, Application US/10052817  
; Publication No. US20020114792A1  
; GENERAL INFORMATION:  
; APPLICANT: Tanzi, Rudolph E.  
; APPLICANT: Kovacs, Dora  
; APPLICANT: Saunders, Aleister J.  
; TITLE OF INVENTION: Alpha-2-Macroglobulin Therapies and Drug Screening Methods for  
; FILE REFERENCE: 0609, 4460005  
; CURRENT APPLICATION NUMBER: US/10/052, 817  
; CURRENT FILING DATE: 2002-01-23

```

; PRIOR APPLICATION NUMBER: 09/241,606
; PRIOR FILING DATE: 1999-02-02
; PRIOR APPLICATION NUMBER: 09/148,503
; PRIOR FILING DATE: 1998-09-04
; PRIOR APPLICATION NUMBER: 60/093,297
; PRIOR FILING DATE: 1998-07-17
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: Patent Ver. 2.0
; SEQ ID NO 2
; LENGTH: 1474
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-052-817-2

```

```

Query Match      19.6%; Score 1441; DB 14; Length 1474;
Best Local Similarity 29.1%; Pred. No. 5.5e-109;
Matches 451; Conservative 266; Mismatches 551; Indels 284; Gaps 59;

```

```

QY 12 LILCVTAALAAVAPREFVLTAPGIIIRPGNVITIGVELLEHSCPSQVTVKALIKTASNLTV 71
DB 15 LILVLPDASVSGKQYQWVLVPSLHHT-ETTEKGCVLISYLNELVTVASLESVRGNSL 73
QY 72 -SVLAEG-VPEKSGKTLTLPISPLNSADE--IYELAVTGRTODELLFSNSTRLSPET 126
DB 74 FTDLERENDVLHCVAF-----AVPKSSNEEMVFLVQKGPTE--FKKRTVVMYKN 124
QY 127 KRIISVFIOTDKALYKPKQKQKFRIVTLFSDPKPYKTSNLIL-IKOPKSNLIQWMLSQSD 185
DB 125 EBSLVFVQDKSIYKFGQTVKRVVSMDENFHPLMLVLTIODPKRIKQWQSFOLE 184
QY 186 LCVISKTQOLSHPLIGDMSIOVQVND--QTYVQSFQVSEYVLPKREVTLLQPLCYMNS 243
DB 185 GGLKQSFPLSEPFQSGSYKVVQKSGGRTEH-PTVEFVLPRKREVQVTPKTIITILE 243
QY 244 KHLNGITTKATYTKYKPKVDYTLPLPL-----SFMCK 276
DB 244 EEMNVSVCGLYTYKRPVGHVTVSICRKYSDASDCHGEDSOAFCEKFSQOLSHGCFYQO 303
QY 277 KKNITTFKINGSANFSEFDEEMKQVNDSSNGLSSEYLDLSPGVEILT-----TYTBSV 331
DB 304 VK--TKVFOLEK-----KEYEMK-----LHLEAQIOEGIVVELTGROSSSEITRTI 347
QY 332 TGISRNVSTNVEFKQHDYILIEFPDYTYTLKPSLNFATVATVRADGNQTLLEERRNNVY 391
DB 348 TKLS-FVKVDHPRQG--LPPFG-----QVRLVQKGPPI--NKYI 384
QY 392 TYTORNTYIYSGNSGNQKMAVQ-KINYT--VPOSGTFKLEP-----ILEDS 438
DB 385 FI-RGNEANY--SNATTDHGLVQFSINTVWVGSLTVRVNYKDRSPCYGVQWVSEH 441
QY 439 SELQKAYVLSGSSMAVHSLFKSPSKTYIOLKTDENIKVGS---PELVVSGNK--R 492
DB 442 EEAHHTAY-----LVPSBSKSFVHLEPMSHELPCHTQTVQAHYTLNGTLLG 489
QY 493 LKELS--YVWVSRGOLVAVG-----KONST-----MFSLTPENSWTPKACVITYIEDDG 540
DB 490 LKKTSTYVILMAKGGIVRTGTHGLVQKQDMKGPISISIPVSDIAPARALLIYAVLPFG 549
QY 541 EIIISVLTLPVQLVFNKIKIKLYMSKYKAPSPKSVLSRISTOPDSIVIGIIVADKSVNLMN 600
DB 550 DVIQSAKDVENCILANKVDLSFSPQSILPASHAHLRYAA--PQVICALRAVDQSLMK 608
QY 601 ASNDITMEVNHLEL-ELVNTGY-----YLGMFNNS-----FA 631
DB 609 PDAEISASSVYVNLPEPKDLTGFPPLNDODEDCINRHHVYINGITTYTPVSSSTNEKQMS 668
QY 632 VFOECGLAWLTDLANTLKDYIDGVYDAA-----YARFMEENEGHIVLDIHDFSL 680
DB 669 FLEDMGLKAFNTSKIRKPKMCPQLOQYEMHGPBGLRVGVEEDVWGRGAKRLVHYBE--- 725
QY 681 GSSPH--VRKHPEETWILDTNMGRYIOEFEVTVPSITSWVATGVFISDLGLT 737
DB 726 ---PHTEVTKYFPETWILDLVWVNSAGVAVGVTVPTDITEMKAGAFCLSDAGIGISS 782

```

```

QY 738 TPVELQAFQPFPIPLNTPYSVIRGEEFALETITFNLYKDATEVXVITIEKSDKPDILMTSS 797
DB 763 T-ASLARQPFVVELTWPYSVIRGBAFLTKATVNLVLPKICRVSGVQLEASPAFLAVPEK 841
QY 798 E-----INATHQOQLVPSSEDGATVLPPIPTHGELPIPTVYALS-----PTA 841
DB 842 EAPHCICANGRO-----TVSWAVPKSIGVNFVSAEALESOELCGTEVPBV 890
QY 842 -----SDAVQMLLYKAGIEKYSQSLDLNLRLOSTLTKTISFSPPTVTVSGEVQ 896
DB 891 PEHGRKDTVIXPLVPEBGELEKETTFNSLLCPSGGEYS--EELSCLKPWNVEBSABAS 947
QY 897 ITAIGDVLGPSINGLASIRMPYCGCEQNMINFANITYIIDLTYKKQTLNLEKALSFP 956
DB 948 VSVIGDILGSAQMONTQNLQMPYCGCEQNMVLPANIVLDYLMETQQLTEIKSKAICY 1007
QY 957 MRQGOEELLYQREDGFSFAG-NYDS9-GSTWLSAPVLRCEFLADPYIIDQVLAHRTY 1014
DB 1008 LNTGYORQLNXYKHVDGYSYTFGERYGRNQNTWLTAFVLKTFQAQARAVITFIDEAHITQAL 1067
QY 1015 TWLKGHOKSNEBFDPGVHISELOGKNSPVTLTAVISLGGYRKQPNIDQESIH 1074
DB 1068 TWLSQRQKNCFCRSGSLNNNAIKGVEDEVTLISAVITTLALBEIPLTVTHPVVRNALFC 1127
QY 1075 LES-----EFSRGISDNYTLALITYALSVSQS-PKAKEALNMTWRAOEGQWFM--- 1124
DB 1128 LESAMKTQOEBDHG-SHYTTQALALAYAFALAGNDKREVKLSINEBAVKQNSVHMERP 1186
QY 1125 VSSBSKSDSWQPR-SLIDIEVAAYALLSHFLQO--TSBGP--IMKLSRQRNSTL 1175
DB 1187 QKRAPVGHFYEPQPSAEVEMTSYVLLA-YLTQAPAPTSBDLTSATNIVKMTQOQAQ 1245
QY 1176 GGFASOTTYTALALSFALNMTERTNIQVYTYGSSBPLAVQO----- 1222
DB 1246 GGFSTQDTVVALHSLKYGATFT-RTGKAQVTTIOSGTFSSPFQVNNRLLLOQS 1304
QY 1223 ---PMAVNIANGFGFALCOLNVVNVKAGSSRRRSIONOBAFDLDVAVK--ENKD 1275
DB 1305 LPELPEBYSMTYBGCYVLOTLSTKNI-----LEKEEFPALGVQTLPTQCD 1353
QY 1276 DLN-HVDLNVCTSPSGP--RSGMALMEVNLGFWPSEASISL--SETVKKVEYDHG 1328
DB 1354 EPKATHSFQISLSTSYGSRASAMAIYDVWVGSGFIPLKPTVWMLERSNHSRTEVSSN 1413
QY 1329 KNLVLSVNETQFCVNIIPAVNRKVSNTQDASISYVDYEPKQOAVASYS 1380
DB 1414 HVLVYLDKVSQTLSPFTVLDVVDVLDKPAIVKYVDYETDEFALAEVNA 1465

```

```

RESULT 6
US-09-873-403-5
; Sequence 5, Application US/09873403
; Patent No. US20020028207A1
; GENERAL INFORMATION:
; APPLICANT: Sistrava, Pramod K
; TITLE OF INVENTION: COMPLEXES OF ALPHA (2) MACROGLOBULIN AND ANTIGENIC
; FILE OF INVENTION: MOLECULES FOR IMMUNOTHERAPY
; CURRENT APPLICATION NUMBER: US/09/873,403
; CURRENT FILING DATE: 2001-06-04
; PRIOR APPLICATION NUMBER: 09/625,139
; PRIOR FILING DATE: 2000-07-25
; PRIOR APPLICATION NUMBER: 60/209,266
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: Fastrsq for Windows Version 3.0
; SEQ ID NO 5
; LENGTH: 1474
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-873-403-5

```

Query Match 19.6%; Score 1440; DB 9; Length 1474;  
 Best Local Similarity 29.1%; Pred. No. 6.7e-109;  
 Matches 451; Conservative 266; Mismatches 551; Indels 284; Gaps 59;

12 LVCCTAALAVAPGPFLLVTAAGIIRPGGNTTIGVLEHCPQVTVKAEELTKASNLTV 71  
 15 LVLVLPDASVSGKQYVNLVPSLHT-ETTEKGCVLSTYLNETVYASLESVGRNLSL 73  
 72 -SVLEABG-VFEKSGFKTLTLPSPLENSADE---IYELRYGTODEILFNSNTRLSFET 126  
 74 FTDLEANDVLHCVAF-----AVPKSSNEEWMLTVQVGPQDE---FKKRTTVWVKN 124  
 127 KRISVFIOTKALYKPKQEVKFRIVTLFSDPKPKYKTSNLIL-IDPKSNLIQMLSQSD 185  
 125 EDSLVFVQTDKSIYKPGQTVKFRVYVMDENFHLNELPLVYIQDPKRNIAQMSFQLE 184  
 186 LGVISTKQFSLSHPLTGDMSIQVQVND--QTYQSGFQVSEVVLKPEFVTLQTPLYCSMNS 243  
 185 GGLQKQFSPSSSEPPQSGYKVVQKSGGRTGH-PFTVEFVLKPEVQVTVPKITITLE 243  
 244 KHLNGTIFAKTYGKPVKGDVTLTFLPL-----SFMGK 276  
 244 EEMVSVGGLTYTGKPGVGHVTVSICRKYSDASDCHGSDSOAFCEKPSGQLNSHGCPTQ 303  
 277 KKNITKTFKINGSANFSFNDKMNWDSNGLSEYLDLSSPGVEYILT---TVTESV 331  
 304 VK--TKVFOAKR-----KEYEMK-----LHTEAQIOEGTVVELTGRSSEITRTI 347  
 332 TGISRNSTNVFKQKHDIIEFPDVTYKXSLNFTAVKTRADGNOLIEERRNNVYI 391  
 348 TKLS-FYVVDHFRQG---IPFFG-----QVRLVDGKGVPIP---NKVI 384  
 392 TVTQRNTEYWSGNSGQKMEAVQ-KINYT--VPOSGTFKIEPP-----ILEDS 438  
 385 FI-KGNENANY--SNATTDHGLVQFSINTNNWCTSLTVAVNKKDSPCGYQWVSEBH 441  
 439 SELQKAYFLGSKSMVAHSLFKSPSKTYIQLKTRDENIKVGS---PFEVLSGNK--R 492  
 442 EEAHTAY-----LVFSPSKSFVHLBPMHSHLPCGHITQVQAHYLLGCTLLG 489  
 493 LKEHS--YMVVSRQULVAVG-----KONST---MPSLTBENSWTPKACIVYIYEDG 540  
 490 LKKLSFYLLIMAKGIVRTGHLGLLVKQEDMKHFSISIPKSIDAPARLLIYAVLPTG 549  
 541 EISDVLIKIPVOLFKNKIKLYMSKVKAEPSEKYSILSYVOPPSIYQIVAVDSKVLNM 600  
 550 DVIQDSAKYDVNCLANKVDSLSPSSQSLPASHAHLKTA--POSCALRAVDOSVLMM 608  
 601 ASNDITMENVHEL-ELYNTGY-----YLGEMFMS-----FA 631  
 609 PDAELASASVYNLPEKDLTGFPGLINDODECINRNINVIINGITYTPVASTNEKOMYS 668  
 632 VPOEGGLVLTDAULTKYIDGYIDNAE-----YARMEEHEGIIYVINDPSL 680  
 669 FLEWGLKAFNNSKIRKPKMCPOLQOQYEMHGPBGLRVGFYSDVDWGRHARLVHER--- 725  
 681 GSSPH---VRHGPETWIMDLTNMGYRYOEFEVTVPDPSISWATGVFVISEDGLGLTT 737  
 726 ---PHTETVKKTFETWIMDLVVVNSAGVAVTVPPTITEMKAGAFCSIEDGLGITS 782  
 738 TPVELQAFQPFITFLNLPYSVIRGEFALETITFNYLKDATERVYVILIEKSDKFDILMTSS 797  
 783 T-ASLRARQPFVELTMYSVIRGEAFTLKATVNLKPKIRVGVQLEASAPFLAVPEK 841  
 798 E-----INATHQOGLLVSESDGATVLPFIRPMLGELPIYVYLS-----PTA 841  
 842 EADPHCICANGRO-----TVSMAVTPKSIKGNVNFVSABALSSOELCTEVPVSV 890  
 842 ---SDAVTMILYKAGIEKYSQSILDLTLNRLQSTLKTLSFSFPPTVYVSGSRVQ 896  
 891 PEHRKQDVLPPLVEBPGLEKETTFNSLLCPSSGGEVS--EELSLKLPVNVESASAS 947  
 897 ITAIGDVLGPIINGLASLIRMPYCGGEQNMINFANIIYLDYLTKKQLTDLNLEKALS 956

12 LVCCTAALAVAPGPFLLVTAAGIIRPGGNTTIGVLEHCPQVTVKAEELTKASNLTV 71  
 15 LVLVLPDASVSGKQYVNLVPSLHT-ETTEKGCVLSTYLNETVYASLESVGRNLSL 73  
 72 -SVLEABG-VFEKSGFKTLTLPSPLENSADE---IYELRYGTODEILFNSNTRLSFET 126  
 74 FTDLEANDVLHCVAF-----AVPKSSNEEWMLTVQVGPQDE---FKKRTTVWVKN 124  
 127 KRISVFIOTKALYKPKQEVKFRIVTLFSDPKPKYKTSNLIL-IDPKSNLIQMLSQSD 185  
 125 EDSLVFVQTDKSIYKPGQTVKFRVYVMDENFHLNELPLVYIQDPKRNIAQMSFQLE 184  
 186 LGVISTKQFSLSHPLTGDMSIQVQVND--QTYQSGFQVSEVVLKPEFVTLQTPLYCSMNS 243  
 185 GGLQKQFSPSSSEPPQSGYKVVQKSGGRTGH-PFTVEFVLKPEVQVTVPKITITLE 243  
 244 KHLNGTIFAKTYGKPVKGDVTLTFLPL-----SFMGK 276  
 244 EEMVSVGGLTYTGKPGVGHVTVSICRKYSDASDCHGSDSOAFCEKPSGQLNSHGCPTQ 303  
 277 KKNITKTFKINGSANFSFNDKMNWDSNGLSEYLDLSSPGVEYILT---TVTESV 331  
 304 VK--TKVFOAKR-----KEYEMK-----LHTEAQIOEGTVVELTGRSSEITRTI 347  
 332 TGISRNSTNVFKQKHDIIEFPDVTYKXSLNFTAVKTRADGNOLIEERRNNVYI 391  
 348 TKLS-FYVVDHFRQG---IPFFG-----QVRLVDGKGVPIP---NKVI 384  
 392 TVTQRNTEYWSGNSGQKMEAVQ-KINYT--VPOSGTFKIEPP-----ILEDS 438  
 385 FI-KGNENANY--SNATTDHGLVQFSINTNNWCTSLTVAVNKKDSPCGYQWVSEBH 441  
 439 SELQKAYFLGSKSMVAHSLFKSPSKTYIQLKTRDENIKVGS---PFEVLSGNK--R 492  
 442 EEAHTAY-----LVFSPSKSFVHLBPMHSHLPCGHITQVQAHYLLGCTLLG 489  
 493 LKEHS--YMVVSRQULVAVG-----KONST---MPSLTBENSWTPKACIVYIYEDG 540  
 490 LKKLSFYLLIMAKGIVRTGHLGLLVKQEDMKHFSISIPKSIDAPARLLIYAVLPTG 549  
 541 EISDVLIKIPVOLFKNKIKLYMSKVKAEPSEKYSILSYVOPPSIYQIVAVDSKVLNM 600  
 550 DVIQDSAKYDVNCLANKVDSLSPSSQSLPASHAHLKTA--POSCALRAVDOSVLMM 608  
 601 ASNDITMENVHEL-ELYNTGY-----YLGEMFMS-----FA 631  
 609 PDAELASASVYNLPEKDLTGFPGLINDODECINRNINVIINGITYTPVASTNEKOMYS 668  
 632 VPOEGGLVLTDAULTKYIDGYIDNAE-----YARMEEHEGIIYVINDPSL 680  
 669 FLEWGLKAFNNSKIRKPKMCPOLQOQYEMHGPBGLRVGFYSDVDWGRHARLVHER--- 725  
 681 GSSPH---VRHGPETWIMDLTNMGYRYOEFEVTVPDPSISWATGVFVISEDGLGLTT 737  
 726 ---PHTETVKKTFETWIMDLVVVNSAGVAVTVPPTITEMKAGAFCSIEDGLGITS 782  
 738 TPVELQAFQPFITFLNLPYSVIRGEFALETITFNYLKDATERVYVILIEKSDKFDILMTSS 797  
 783 T-ASLRARQPFVELTMYSVIRGEAFTLKATVNLKPKIRVGVQLEASAPFLAVPEK 841  
 798 E-----INATHQOGLLVSESDGATVLPFIRPMLGELPIYVYLS-----PTA 841  
 842 EADPHCICANGRO-----TVSMAVTPKSIKGNVNFVSABALSSOELCTEVPVSV 890  
 842 ---SDAVTMILYKAGIEKYSQSILDLTLNRLQSTLKTLSFSFPPTVYVSGSRVQ 896  
 891 PEHRKQDVLPPLVEBPGLEKETTFNSLLCPSSGGEVS--EELSLKLPVNVESASAS 947  
 897 ITAIGDVLGPIINGLASLIRMPYCGGEQNMINFANIIYLDYLTKKQLTDLNLEKALS 956

US-10-292-081A-10  
 ; Sequence 10, Application US/10292081A  
 ; Publication No. US20030162202A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Kenneth David Becker  
 ; APPLICANT: Gonul Velliclebi  
 ; APPLICANT: Xin Wang  
 ; APPLICANT: Randolph E. Tanzi  
 ; APPLICANT: Lars Bertram  
 ; APPLICANT: Aleister J. Saunders  
 ; TITLE OF INVENTION: SINGLE NUCLEOTIDE POLYMORPHISMS AND MUTATIONS ON ALPHA-2-MACROGLOB  
 ; FILE REFERENCE: 37481-3323  
 ; CURRENT APPLICATION NUMBER: US/10/292, 081A  
 ; CURRENT FILING DATE: 2002-11-08  
 ; PRIOR APPLICATION NUMBER: 60/337434  
 ; PRIOR FILING DATE: 2001-11-09  
 ; NUMBER OF SEQ ID NOS: 15  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 10  
 ; LENGTH: 1474  
 ; TYPE: PRN  
 ; ORGANISM: Homo sapiens  
 ; US-10-292-081A-10

Query Match 19.6%; Score 1440; DB 12; Length 1474;  
 Best Local Similarity 29.1%; Pred. No. 6.7e-109;  
 Matches 451; Conservative 266; Mismatches 551; Indels 284; Gaps 59;

12 LVCCTAALAVAPGPFLLVTAAGIIRPGGNTTIGVLEHCPQVTVKAEELTKASNLTV 71  
 15 LVLVLPDASVSGKQYVNLVPSLHT-ETTEKGCVLSTYLNETVYASLESVGRNLSL 73  
 72 -SVLEABG-VFEKSGFKTLTLPSPLENSADE---IYELRYGTODEILFNSNTRLSFET 126  
 74 FTDLEANDVLHCVAF-----AVPKSSNEEWMLTVQVGPQDE---FKKRTTVWVKN 124

```

OY 127 KRISFIOTDKALYKQEVKFRITVTLFSDFKPYKTSNLIL- IKOPKSNLIOQMLSOQSD 185
DB 125 EDSLVEFVQDLSIKYRGQTVKFRVVSMDENFHPINELPLVYIOPDKNRIAQWOSFOLE 184
OY 186 LGVISTKFOLSHPLIDGMSIOVOND--QTYOSFOVSEYVLPKFEVTLQTPLYCSMNS 243
DB 185 GGLKQFSPFLSSEPPQSGSYKVVQKSGGRTEH-PTVEEVLVLPKFEVQVTPKTIITILE 243
OY 244 KHLNGTITAKTYTKPKVGDVTLTFLPL-----SFMGK 276
DB 244 EEMNVSVGLTYTKPKVPGHVTVSI CRKYSDASDCHGEDSOAFCEKFSQQLNSHGCFFQO 303
OY 277 KKNITKTKINGSANFSFDEEMKQVMSNGLSEYLDLSRPPYEILT-----TVTESV 331
DB 304 VK--TKVFKLR-----KEYEMK-----LHTEAQIOEGTVVETLGRQSEITRTI 347
OY 332 TGISNVSTNVFPGKHDTIEFPDVTYVLPKPSLNTATVYKVRAGNOQTLTEERANNVI 391
DB 348 TKLS-FVKVDSHFQO--IPFG-----QVRLVQKGVPIP---NKVI 384
OY 392 TVTQNTLEYWGSNSGNOKEAVO-KINYT--VFQSGTFKIEFP-----ILEDSS 438
DB 385 FI-RGNENANY--SNATTDHGLVQFSINTNMGTSLTVRNVYKDRSPCYGQWSEBH 441
OY 439 SELQKAVFLGSKSSMAVHSLFKSPSKTYIOLKTRDENIKVGS---PEELVSGNK--R 492
DB 442 EEAHHTAY-----LVFSPSKSFVHLEPMSHELPGHQTQVQAHYILNGSTLLG 489
OY 493 LKELS--YVVSRGOLVAVG-----KONST---MFSITPENSMTPPACVITVYIEDDG 540
DB 490 LKQSFYVLIWAKGGIVTGHGLLVKQEDMKHRSISIPVKSADLAPARLITVAVLPTG 549
OY 541 EHSIVLKIPIVQVLFKNKIKLYWASKVABPSEKVSLSISVOPDSIVGVAVDKSVNLMN 600
DB 550 DVIGSADYDENCANAKVDSFSPSOSLPASHALHRTAA--POSVCALRAVDOSVLAK 608
OY 601 ASNDITMENVHEL-ELYNTGY-----YLGMPFNS-----FA 631
DB 609 PDAELSSASVYVNLLEKOLTFPGPLNQDDEDCINRHNVIINGITYPVSTNEKDMVS 668
OY 632 VFQEGKWLVDANLTKOYIDGVYNAE-----YAEFMEBEGHIVDIHDSL 680
DB 669 FLEDGLKAFNTSKRKPKMCPOLQOYEWGPEGLRVGFESDVGARGHARLVHEB--- 725
OY 681 GSSPH---VRKHEPTWIMDTNMGYRTYQEFVYVUPDSISVAVTNGVFISDLGLTY 737
DB 726 ---PTEETVRKTFPEBTWIMDLVVVNSAGVAVGVTVPDITTEWKAGACLSLSDAGLSS 782
OY 738 TPVELQATOPPEFPLNLPYSVIRGEFALETIFNYLKDQATEVKYIIEKDFDIAMTSS 797
DB 783 T-ASIRARQFPFVELTWPVSIRGEAFILKATVANYLPRCIRVSQQLASPAFLAVPREK 841
OY 798 E-----INATGHQOTLLVPSSEDGATVLPPIRPTHGEIPIVYALS-----PTA 841
DB 842 EQAPHICANGRO-----TVSWAVTPKSLGAVNFTVSAEALSSOELCGTEVPSV 890
OY 842 -----SDAVQMLIVKAGIEKSYQSSTILLDTNRLOSTKLTSFSPRPNTVTSERVQ 896
DB 891 PEHGRKDTVIKPLVPEGLEXETTFNSLCPSSGEVS---BELSLKLPENVEBSAAS 947
OY 897 ITAIDVLAGPSINGLASTLRMPYGGCEONMIFANITYILDVLTKKOLTDNLKEXALSF 956
DB 948 VSVGLDIIIGSAMONTNLOMPYGGCEONMVFANITYILDVLTNETOOLTEPVSKAIGY 1007
OY 957 MRQGYRELLYQREDGSFAFG-NDPS--GSTWLSAPVLRCELPADPYIDIDQVILHRTY 1014
DB 1008 INTGYQDOLNMGYDGSYTFEGRYGRNGNTWLAFLVKTFAQARAYIFIDEAHTIQA 1067
OY 1015 TWLKHQKSNGBFMPKRVHSELQKNSPYTLTAVYVLSLGRKQOPNDVOESLHF 1074
DB 1068 IWLQRQKQKNGCFRSGSLNNAIKGVEDEVTLTAVYITIALLEPLTVTHPVVNAALFC 1127
OY 1075 LES-----EFGSGISDNVTLLALITYALSVOGS-PEAKEALMMLTWRAEQSGMQFW--- 1124

```

```

DB 1128 LESAKTKAQEDHG-SHYTVALALAYAPALAGNODKREKVKSLNEBAVKDQNSVHERP 1186
OY 1125 VSSSEKSDSMQPR--SIDIEVAVALISHLQOQ---TSRGP-----IMWLSRORSL 1175
DB 1187 OKPKAPVGHFPPQAPSVEKETSIVLA-VLTQOPAPTSIDLTSATNIVKMTKQNAQ 1245
OY 1176 GGFASTODTVALKALSEFALMNTERTINQVTVGSPSPPLAVQ----- 1222
DB 1246 GGFSTODTVALALSLKSGAATF-RIGKAAQVLTIGSGFSSKFOVDNNRLLLOQVS 1304
OY 1223 ---PMAVINSANGFALICOLNVYVNVKASGSSRRRSIONQEAFLDVAK--ENKD 1275
DB 1305 LPELFGESYKMTGEGGCVYIOTSILKYN-----LPEKEEFPALGVQTLPTQCD 1353
OY 1276 DLN-HVDLNVCTSPGCG--RSGALMEVNLISGFVWPSAISL--SEVYKVEYDHG 1328
DB 1354 EPKATSFQISLSVYSYTSRSASNNALVDVWVSGFPLPKETVWMLERSNHVSRTVESSN 1413
OY 1349 KNLVLDVSNETQFCVNIPIAVNPKVSNTOASVSIYDYEPBROAVRSYNS 1380
DB 1414 HVLITLDKVSQTLSTLFPFTVQDVPRDLKRAIVKVDYETDEFAIETNA 1465

```

```

RESULT 8
US-10-292-081A-12
; Sequence 12, Application US/10292081A
; Publication No. US20030162202A1
; GENERAL INFORMATION:
; APPLICANT: Kenneth David Becker
; APPLICANT: Gonul Velicetlebi
; APPLICANT: Xin Wang
; APPLICANT: Randolph E. Tanzi
; APPLICANT: Lars Bertam
; APPLICANT: Aleister J. Saunders
; TITLE OF INVENTION: SINGLE NUCLEOTIDE POLYMORPHISMS AND MUTATIONS ON ALPHA-2-MACROGLO-
; FILE REFERENCE: 37481-3323
; CURRENT APPLICATION NUMBER: US/10/292,081A
; PRIOR FILING DATE: 2002-11-08
; PRIOR APPLICATION NUMBER: 60/337434
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12
; LENGTH: 1474
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-292-081A-12

```

```

Query Match 19.6%; Score 1440; DB 12; Length 1474;
Best Local Similarity 29.1%; Pred. No. 6.7e-109;
Matches 451; Conservative 266; Mismatches 551; Indels 284; Gaps 59;
OY 12 LUCVCTALAAVAPREFVLTAPGIIIRPGANTIGVELLEHCPQVTVAKELTASNLTV 71
DB 15 LVLLPDPASVSGRQVWLVPSLHT-ETTEKGCVLISYANETVYSASLESVRGRSL 73
OY 72 -SVLABG-VFEKGSFKTLTLPSPLSADE---IYELRVGTQODELLFNSSTRLSFET 126
DB 74 FTDLAENDVLAHCAF-----AVKSSSNEVWFVQVGPQOE---FKKRTVWVKN 124
OY 127 KRISFIOTDKALYKQEVKFRITVTLFSDFKPYKTSNLIL- IKOPKSNLIOQMLSOQSD 185
DB 125 EDSLVEFVQDLSIKYRGQTVKFRVVSMDENFHPINELPLVYIOPDKNRIAQWOSFOLE 184
OY 186 LGVISTKFOLSHPLIDGMSIOVOND--QTYOSFOVSEYVLPKFEVTLQTPLYCSMNS 243
DB 185 GGLKQFSPFLSSEPPQSGSYKVVQKSGGRTEH-PTVEEVLVLPKFEVQVTPKTIITILE 243
OY 244 KHLNGTITAKTYTKPKVGDVTLTFLPL-----SFMGK 276
DB 244 EEMNVSVGLTYTKPKVPGHVTVSI CRKYSDASDCHGEDSOAFCEKFSQQLNSHGCFFQO 303

```

```

QY 277 KKNITTKFKINGSANFSFNDEEMKNVMSNGSLSEYLDLSSPPVEIIT-----TVESV 331
Db 304 VK--TKVFOLEK-----KEYEMK-----LHTEAQIOEBGTVEITLGRSSEITRTI 347
QY 332 TGISRNVSTNVFKOHDIIEFPDYTTVLKPSLNFATVKTTRADGNQLTLEBRNNVVI 391
Db 348 TKLS-FKAVDSHFRQG---IPFG-----QRLVDGKGVPIP---NKVI 384
QY 392 TVTGRNTEYWSGNSGNQMEAVQ-KINYT--VPOSGTFKIEFP-----ILED 438
Db 385 FI-RGNEANY--SNATTDHGLVQFSINTTNWGTSLTVRVNKKDRSPCYQWVSEEH 441
QY 439 SELQKAYFLGSKSMVAHSLFKSPSKTYIOLKTRDENIKVGS---PFEIVSGNK--R 492
Db 442 EEAHTAY-----LVFSPSKFVHLEPMSHELPCGHQTQVQAHYIINGGL 489
QY 493 LKELS--YMWVSROGLVAVG-----KONST---MPSLTENSWTPKACIVYIIEBDG 540
Db 490 LKKSIFYLLMAKGIYVGTGHLVQEDMKHPSISIPKSDIAPARLLIYAVLPTG 549
QY 541 EISDVLKIPVQVFNKNIKLYWSKVKAPESEKVSLSISVTPQDSIVGIVADKSVMILN 600
Db 550 DVIGDSAKYDVENCANKVDSLSPSSQSLPASHMLHRTA--PQVCALRAVDQSVLLMK 608
QY 601 ASNDITMENVHEL-ELYNTGY-----YIGMFMS-----FA 631
Db 609 PDAELSSAVYNLPEKDLTGFPGLNDQDEDCINRHNVIYNGITYTPVSTNEKDMYS 668
QY 632 VFOEGMLVLDANLTKYIDGVYDNE-----YARMEENEGHIVIHPSL 680
Db 669 FLEWGLAFNNSKIRKPKCPQLOQYEMHGPGLRVGFYSDDVGRCHARLVHEE--- 725
QY 681 GSSPH--VRKHFPETWIMLDTNMGRYIOEFEVTPPDSISWVATGVISDGLGLTT 737
Db 726 ---HTETVRYKFPETWIMDLVNVASAGVAGVTPPTTEKAGAFCLSEDGLGIS 782
QY 738 TPVELQAFQPFIFLNLPSYVIRGEFALITTFNYLKDATEVYVIEKSDKFDILMTSS 797
Db 783 T-ASIRAPQPFVELTWPYSYIRGEAFILKATVNLTKCIRVSVQLEASAPFLAVPEK 841
QY 798 E-----INATGHOQLVPSDEGATVLPFIRPHLGEIPIVYALS-----PTA 841
Db 842 EDAHPCICANGRO-----TVSWAVTPKSLGNVNFVSAEALSSOELCSTEVPSV 890
QY 842 ----SDAVTOMILYKAGEIKESYSOSITLDTNRLDOSTLKTISFSPPTVATGSEVQ 896
Db 891 PEHGRKDVYKPLVEPBGLEKETTFNSLLCPGSEVS---EELSLKLPNVVESARAS 947
QY 897 ITAIGDVLGPSINGLASLIRMPYGGGEQNMIFAPNIYILDYLTKKOQLTDLNEKALSF 956
Db 948 VSVLGDILGAMQNTQNLQMPYGGGEQNMVLPAPNIYLDYLNBTQQLTREVSKAIGY 1007
QY 957 MRQGYQRELLYQREDEGSFAG-NYDES-GSTWLSAFYLRCPLEADPYIDIDQVLRHTY 1014
Db 1008 LNTGFORQNTKYHDSYSTGERYGRQGTWMLTAFILKTFQAQRAVYIFIDEAHITQAL 1067
QY 1015 TWLKGHOSNGEPMDPGRVHISELQGNKSPVTLTAVYVSLGGRKQPINIDVQESIH 1074
Db 1068 TWLSORQKQNGCFNSSGSLNNNAIKGVEDEVTLTASVYITALLPLTTHTPVRNALPFC 1127
QY 1075 LES-----EFSRGISDNYTLLALITYALSVS--PKAKEALNMLTWAEQGGQFW-- 1124
Db 1128 LESAMKTAQEBDGH-SHYTTKALLAYAPALAGNDKREVKLSINEBAVKDKNVHERP 1186
QY 1125 VSSSKLSDSQWR--SLDIEVAVALISHFLQFO--TSEGIP---IKMILSRQNSL 1175
Db 1187 QKPRAPVGHFEYPOAPSALVEMTSYVILA-YLTQAPATSSEDLSATVIVKMITKQNAQ 1245
QY 1176 GGFASDTOTVALKALSEFALNMTERTNIQVTVGPSSPPLAVVQ----- 1222
Db 1246 GGFSTSDTVALHALSKYGAATFL-RTKAQAQVYITQSSGTFSSKFOVDNNRLLQOVS 1304
QY 1223 ----PMAVNISANGFPAICQNLNVYVNVKAGSSRRRSIQNGEAFDLDAVK--ENKD 1275

```

```

Db 1305 LPELPEYSMKYVGEVGVYQTLISLKNYI-----LPEKEEPPFALGVQTLQPTCD 1353
QY 1276 DLN-HVDLANCTSESGRG--RSGMALMEYNLLSGFVPEALSL--SETVKVEVDHG 1328
Db 1354 EPKAHTSFQISLSVSTYSGSASNMALVDYVMSGFLPKPYTKMLERSHVSTREYSSN 1413
QY 1329 KNLVYDSVNETQFCVNIPAVRNFKVSNTOGASVSIYDVEPRROAVRSYNS 1380
Db 1414 HVLIVDKVSNQTLSEFTVLDQVPRDLKPAIVKYDYETDEPAIAEYNA 1465

RESULT 9
US-10-292-081A-13
; Sequence 13, Application US/10292081A
; Publication No. US20030162202A1
; GENERAL INFORMATION:
; APPLICANT: Kenneth David Becker
; APPLICANT: Genul Velicelabi
; APPLICANT: Xin Wang
; APPLICANT: Randolph E. Tanzi
; APPLICANT: Lars Bertram
; APPLICANT: Aleister J. Saunders
; TITLE OR INVENTION: SINGLE NUCLEOTIDE POLYMORPHISMS AND MUTATIONS ON ALPHA-2-MACROGLOB
; FILE REFERENCE: 37481-3323
; CURRENT APPLICATION NUMBER: US/10/292,081A
; CURRENT FILING DATE: 2002-11-08
; PRIOR APPLICATION NUMBER: 60/337434
; PRIOR FILING DATE: 2001-11-09
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13
; LENGTH: 1474
; TYPE: PR
; ORGANISM: Homo sapiens
US-10-292-081A-13

Query Match 19.6%; Score 1440; DB 12; Length 1474;
Best Local Similarity 29.1%; Pred. No. 6,7e-109;
Matches 451; Conservative 266; Mismatches 551; Indels 284; Gaps 59;

QY 12 LCVCCTAALAVADGPRFLVTAAGIIRPGNVITGVLEHCPQVYKAEILKTASNLTV 71
Db 15 LVLVLPDASVSGKQPMVLPVSLHT-ETTEKGVLLSYLNTVYVSALESVGRNRL 73
QY 72 -SVLEAEG-VFREGSFRTLTPSLPLNSADE---IYELRPTGTODEILFNSNTRLSFET 126
Db 74 FTDLAENDVLACVAF-----AVPKSSNEEVMFLTVQVKGPTQD---FKKRTVWVKN 124
QY 127 KRISVFIQTDKALYKPKQEVKFRITVLPDFPKPYKTSNLTL-IKDPKSNLIQOMLSQSD 185
Db 125 EDSLVEFQIDKSIYKKGQYKFRVNSMDENFHLNELIFLYIODPKGNRIAGMOSFOLE 184
QY 186 LGVISTKFOULSHPIIGDMSIOVOND--QTYQSFQVSEYVLPKEVTLQTPLYCSMNS 243
Db 185 GGLKQSFPLSSPFGSVYVWVQKXSGGRTEN-PTVEEFVLPKFEVQTVVKIITILE 243
QY 244 KHLNGITTAITYTKGYKGVDTYTLFPL-----SFMK 276
Db 244 EEMNVSVCGLYTGKVPVGHVTVSICRKYSADSDCHGEDSQAPCEKFRSGQLNSHGCFYQ 303
QY 277 KKNITTKFKINGSANFSFNDEEMKNVMSNGSLSEYLDLSSPPVEIIT-----TVESV 331
Db 304 VK--TKVFOLEK-----KEYEMK-----LHTEAQIOEBGTVEITLGRSSEITRTI 347
QY 332 TGISRNVSTNVFKOHDIIEFPDYTTVLKPSLNFATVKTTRADGNQLTLEBRNNVVI 391
Db 348 TKLS-FKAVDSHFRQG---IPFG-----QRLVDGKGVPIP---NKVI 384
QY 392 TVTGRNTEYWSGNSGNQMEAVQ-KINYT--VPOSGTFKIEFP-----ILED 438
Db 385 FI-RGNEANY--SNATTDHGLVQFSINTTNWGTSLTVRVNKKDRSPCYQWVSEEH 441

```





```

Db 348 TKLS-FVVDHFRG--IPFG-----QVRLVDGKVPF--NKVI 384
Qy 392 TVTQNTREYMSGNSGNQKMAVO-KINTY--VPOSGTFRIBP-----LIEDS 438
Db 385 FI-RGENANY--SNATTDEHGVQFSINTTNVMTSLTVAVNATKDSPCYGVMBEER 441
Qy 439 SELQKAVFLGSKSMASHLFFKSPSKTYIOLKTRDENIKVS--PFLVBSGNK--R 492
Db 442 BEAHHTV-----LVFSPSKSFVHLEPMHSHLPCHGHTQYQAHYILNGTLLG 489
Qy 493 LKEIS--YMVVSRQOLVAVG-----KONST--MESLTPENSMTPKACVITYIEDDG 540
Db 490 LKKSIFYLIMAKGIVTGTGHLVQEDMKHFSISIPIKSDIAPARLLIYAVLPTG 549
Qy 541 EIIDVAKIYQVLFKXKIKLYMSKVAEPEKXISIRISVQPSIYQIVAVDKSVMN 600
Db 550 DVIDSAYDVENCILANKVDSFSPSOLPASHAHLRVTAA--PQVCALRAVDOSVLLMK 608
Qy 601 ASNDITMENVHEL-ELYNTGY-----YLGFMFNS-----FA 631
Db 609 PDALSSSVYNNLPEKDLTGFPGRLNDQDDEDCINRNNVINGITPVSTNEXOMYS 668
Qy 632 VFQCGMLVTDANLTQYIDGYDNE-----YERMEENEGHIVIDHPSL 680
Db 669 FLEDMGLKAFNSKIRKPCPOLQOYEMHGPGLRVGFESDYMGRGHARLVABE-- 725
Qy 681 GSSHH---VRKHFPETWIMDTNNGRYOFEFTVPDSTISWATGVSIEDGLT 737
Db 726 ---PHETVRKYFPEWTIMDLVVVNSAGVAVGTVPDITTEMKAGAFCSIEDGLT 782
Qy 738 TPVELQAFOPFIPLNLPYSVIRGEEFALITIEFYLKDATEVKYIIEKSDKPILTSS 797
Db 783 T-ASLRAPQFPFVELTMYSVIRGEAFILKATVNLKPCIRVSVQLEASAFILAVPEK 841
Qy 798 E-----INATGHQOTLLVPSSEDATVLPPIRPHLGIPITVIALS-----PTA 841
Db 842 EQAPHCICANRO-----TVSWAVTPKISIGNVFTVSAEALESQELCTEVPVS 890
Qy 842 -----SDAVTOMILVKAEGIKESYSOSILDLTNRLOSTLKTISFSPPTVTYGSERVQ 896
Db 891 PEHGRKDTIVIKPLVPEEGELETTFNSLCPSGEVS--BELSLKLPVNVESARAS 947
Qy 897 ITAIGDVLGSPINGLASLIRMPYCGGEOMINFAPIYIIDLTKKKOLTDNLKELASLF 956
Db 948 VSVUGDILGSAQNTQNLQMPYCGGEOMNVLFAPIYIIDLNETQOLTFEVSXKALGY 1007
Qy 957 MRQGYORELQYREDSFSAFG-NYDPS-GSTWLSAFVLRCLFADPYIIDQNVLRHTY 1014
Db 1008 LNTGYORQLNKHVDGYSYTFGERYGRNQGTWLTAFVLTKEFQAARAVIFIDEAHITQAL 1067
Qy 1015 TWLKGHQSNGEFPDPRVHISELOGKNSPYTLATVTSILGRYKOPIDVQESIH 1074
Db 1068 TWLSORQKDCNCFSSGSLNNNAIKGVEDEVTLISAYITLALPIPLTVTHPVVRNALFC 1127
Qy 1075 LES-----EFSRGISDNYTLALITVALSVGS--PKAKEALNMLTWABEGOGMFM-- 1124
Db 1128 LESAMKTAQEGDGH-SHYTALALAYAFALAGNODKKEVLKSLNEAVKXKNSVHBERP 1186
Qy 1125 VSSSEKLSDSWOPR--SLDIEVAAYALSHLQFO--TSEGIP-----IMKWSRQNSL 1175
Db 1187 QKRAPYGHFEYEPQAPSAEVEMTSYVLLA-YLTQAPATSEDLASATNIVIMIKQNAQ 1245
Qy 1176 GGFSTQDTTVALKALSEFALAMTERTINQVITYTGPSSPEPLAVVQ----- 1222
Db 1246 GGFSTQDTTVALKALSEFALAMTERTINQVITYTGPSSPEPLAVVQ----- 1222
Qy 1223 ---PMAVNISANGGFALCOLNVVYVNVKASGSRBRRSIQNOEAFDLDVAVK--ENKD 1275
Db 1305 LPELPGEYSKMYTBEGCVYIOTSLKTYNI-----LPEKEEPFALGVQLPOTCD 1353
Qy 1276 DLN-HVDLNVCTSESGP--RSGMALMEVNLISGFVWVPSBAISL--SETVKKVEYDHG 1328

```

```

Db 1354 EPRAHTSFOISLSVSYTGRSASNAIMAVDKWVSGEPIPKPTVKMLERSNHSRTEVSSN 1413
Qy 1329 KNLVYDLSVNETQFCVNIAPVRFKVSNTODASVSIVDYEPBROAVRSYNS 1380
Db 1414 HVLIVDKNSQTLSEFFTVLQDVPRDLKPAIVKYDIYETDEPFAIAEYNA 1465

RESULT 11
US-09-756-247-4
: Sequence 4, Application US/09756247
: Publication NO. US20030180722A1
: GENERAL INFORMATION:
: APPLICANT: Godbole, Shubhada D
: APPLICANT: Boyle, Bryan J
: APPLICANT: Mize, Nancy K
: APPLICANT: Deng, Cenhua
: APPLICANT: Goodrich, Ryle
: APPLICANT: Arterburn, Matthew C
: APPLICANT: Zhou, Ping
: APPLICANT: Tang, Y. Tom
: APPLICANT: Liu, Chenghua
: APPLICANT: Yeung, George
: APPLICANT: Demarec, Radoje T
: TITLE OF INVENTION: METHODS AND MATERIALS RELATING TO ALPHA-2-MACROGLOBULIN-LIKE POI
: FILE REFERENCE: HYS-31CIP
: CURRENT APPLICATION NUMBER: US/09/756,247
: PRIOR FILING DATE: 2001-01-08
: PRIOR APPLICATION NUMBER: 09/649,167
: PRIOR FILING DATE: 2000-08-23
: PRIOR APPLICATION NUMBER: 09/540,217
: PRIOR FILING DATE: 2000-03-31
: PRIOR APPLICATION NUMBER: 09/684,711
: PRIOR FILING DATE: 2000-10-06
: PRIOR APPLICATION NUMBER: 09/560,875
: PRIOR FILING DATE: 2000-04-27
: PRIOR APPLICATION NUMBER: 09/496,914
: PRIOR FILING DATE: 2000-02-03
: NUMBER OF SEQ ID NOS: 41
: SOFTWARE: PatentIn version 3.0
: SEQ ID NO 4
: LENGTH: 1508
: TYPE: PRT
: ORGANISM: Homo sapiens
: US-09-756-247-4

Query Match 19.6%; Score 1437.5; DB 12; Length 1508;
Best Local Similarity 27.1%; Pred. No. 1,le-108;
Matches 422; Conservative 274; Mismatches 552; Indels 307; Gaps 45;

Qy 20 LAVAPG-----PRFLVAPGILR--PGGNVTIGVELLEHCPGQVTVABELKTASNLTVSV 73
Db 10 LALSPALAEELPVYLVLPRLNLP-----SVQKVCULDSBGSDVAFVTLKTKDTQKL 65
Qy 74 LEAGVEFEKSFETLLTPSLPLNSADEIELRYGRGTODEILFNSSTRLSFETKRISVPI 133
Db 66 LEVSGLKRRHLHGISFLVPPAGTEVATIRSG--VGNNISPEKKKVLIQOGNGTFV 124
Qy 134 QTDKALYKPKQEVKFRIVTLFSDPKYKTSNLTL-IKDPKSNLIQWLSQOSDLGVISKT 192
Db 125 QTDKPLVTPQOQYFPIVITMDSNFVNDKYSWELQDPNSNRILAQMLEVAPROGIVDS 184
Qy 193 FOJLSHPHIGDWSIQVANDQTYQSFQVSEVYLPRFEVTLQRPVLCNSMKHLNGTITA 252
Db 185 FOJLAPEMAGTYV--VAVNEGKTFGTFVSEVYLPKFKVEVPEKELSTVQESFLVKIC 242
Qy 253 KYTGGKPVKGDVTLTLPFLSMGKKKNTTKFKINSANESFNDDEMKVWDSNGLSLEY 312
Db 243 RYTYGKPMGAVGVSYC-----QKAN-----TWYREVEREQULPDKRNLISGQ 285
Qy 313 LD-----LSSP-----GPEVILTVTESVTGISRNVSNNVFPKQHDYIIIEFF 354
Db 286 TDKTGCSAPVDWATPDLIGAVYSHQINIVATVVEEGTGEANATONIIYISPMGSMTFE 345

```





QY 319 -----GPEVLLTVTESVTGSRNVSTNVFKOHDIIEFPDYTVLKPSLN 365  
 DB 280 DMATEFDLIGAYASHQINIVATVEEGTVEANATQIYISPOGMSWTFEDTSMFYHPNP 339  
 QY 366 FPAIVKVTADNOGLTEERRNNVY-----TYVORN-----YTEWGS 405  
 DB 340 FSGKLRVRGHDSFL-----KHLVPLVYGTNGTFNQTLVTDNMGLABFTLETSGMNGT 394  
 QY 406 NSGNQKMAVOKINTVTPSGTFKIE-----PFLDSSELQKAYFLSKSMAMH 457  
 DB 395 D-----VSLBGRFQMEDLYNPEQVPRYQNALMLHRPSTSTRSPLGH 439  
 QY 458 SLFKSPSKTYTDLKTRDENIKVSGPEELVY-----SGMKRLKELSYMVWSRGOLVA 508  
 DB 440 RL-----NGPLKCGQPOEVLVDYIDPADSPGOEI-SFSYVILIGKSLVM 484  
 QY 509 VGRON-----STMSLTPENSWTPKACVIVYIEDDEIISDLKIPVQLVKNK 558  
 DB 485 ESOQKLNKKKGLKASFSLSLFTSRPADPSPVIVAIIPSGGVADKIQFSGMCFDQ 544  
 QY 559 IKLVMSKVAEPSEKSLRISTQPSIVGIYAVDSKVLMAANDITMENVHELIN 618  
 DB 545 VSLGSPSQOLPEAEVLOLOAA-PESLCLALRAVDSEVILLRP-----DRELSN 592  
 QY 619 TGYILGMF----- 626  
 DB 593 RSVY-GMFPWYGHVYQVAEYDQCPVSGMDFPQPLIDPMQGHSSQSIIMRPSFSG 651  
 QY 627 MNSFAVQECGLMVLTDANLTQDYIDVYDNMAEY-----ERFMEENGHIVDHD 677  
 DB 652 TDLFSFRRDVGKILSNNAKKR-VDCHSRSEYSTAMGGGHPAEFESTPLHQE--- 707  
 QY 678 FSLGSSPHKRPETWIMLDNMGRIYQEFVYVYDSTISWATGFISEDLGLGTT 737  
 DB 708 -----DSQVROYFPEWIMLDLPFISGSGEAVHVTYDPAITMKAMSFCTOSRGGLSP 762  
 QY 738 TPEVLOAFQPFPLFNLPSYVIRGEFEALFETIFNLYKQATEKVIIEKSDKDIIMTS 797  
 DB 763 T-VGLTAFKFPFVDLTPYSVVRGSEFRLATIFNLYKQICIRQVDILAKSHYOLSWAD 821  
 QY 798 EINATGHQOTLVPSSEDAVTLFPIRPHLGEIPITVTL-----SPYA 841  
 DB 822 S-----QTSSCICADDAKTHHMITAVLGHINFTISKILDSNEPCGQGFVPQKR 875  
 QY 842 SPAYVQMITVKAEG--IEKSYSOSILDDTNRLOSTLKSFSFPNVTYTSERVQITA 899  
 DB 876 SDTLKRPVLVKEGVLEVETHS-----SLCPKGVASVSLELPVDIVPSTRAYTV 930  
 QY 900 IG-----DVLGSPINGLASLIRMPYGGGEONMIPNPIYILDY 938  
 DB 931 LKQLEIIDSEKRRMEAAKWMDINGTALQNDGLVQMPSGGGEONMVLFAPIIYVLO 990  
 QY 939 LTRKQOLDNLEKALSPRGOYQRELLYQREDSFSAFANYDPGSGTWLSAFVLRCFLE 998  
 DB 991 LERAGLITFIRSRBAVGLFIEIGYQKELAMKHSNGSYSAFERDNGENTLTLFVTKCFQ 1050  
 QY 999 APPYIDIDONVLAHTYTTLKGHOKSGEWFDEGRVHSELQGNKSPVTLTAYITSIG 1058  
 DB 1051 AOKFIFIDPKNIQDALKMAAGNOLPBGCYANVGNLHTAMKGVDEVSILTYVTLALLE 1110  
 QY 1059 YARKYONIVOSIHLFSEFSGISDNTYLLALITVALSVSAPRAKE-ALMMLTMRARQ 1117  
 DB 1111 MKGDVDDPMVSOGLRCLKNS-ATSTTNLYTQALALYIFSLAGBMDIRNILLQOLOQAI 1169  
 QY 1118 EGGMQFWSSSEKLSDS--W-OPRSLDIEVAAYALISH-----FLQFQTSSEGIPIMRWLS 1169  
 DB 1170 SGGSIYWSQKPTSSNASPMSEPAAYDELTAVALAQLTRPSLQKEIAXKATSIYAWLA 1229  
 QY 1170 RQNSLIGSGASTODTVALKALSEFA--ALAMTERNIOVTVT-----GSSSGSEPLAV 1220  
 DB 1230 KOHNAGGSSSTODTVALQALAKATYATYAMSEBEINLVKSTENFORTFNIQSVRLIVF 1289

QY 1221 VQ-----PMAVNISANGFPAICQLVNVYVWKAAGSSRRRRS:IONDEAFDLDAVK--- 1271  
 DB 1280 QODTLPNVPGMYTLBASQGCYVQVQTLRLNLP-----TMMKTFSLSVIEGAR 1340  
 QY 1272 -ENKDDLHVDLNVCTSSSGP-GRSGMALMEVNLISGF--MVPSPALISLSE-TYKKVEYD 1326  
 DB 1341 CEQPTSPSLTLTHTTSYVSGRSSSMAIVEVKMLSGFSPMEGTQOLLQOPLVKEVFE 1400  
 QY 1327 HGKMLYLDVSVNEQFCVNIPAVRNFKVSNTQDASVSLVDYEP 1370  
 DB 1401 TDTLNTYIDELIKNTQTYTTFTISQSVLVTKLPATIKYVDYLP 1444  
 RESULT 14  
 US-10-292-081A-9  
 ; Sequence 9, Application US/10292081A  
 ; Publication No. US20030162202A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Kenneth David Becker  
 ; APPLICANT: Gonul Velicelbeli  
 ; APPLICANT: Xin Wang  
 ; APPLICANT: Randolph E. Tanzi  
 ; APPLICANT: Lars Bertram  
 ; APPLICANT: Aleister J. Saunders  
 ; TITLE OF INVENTION: SINGLE NUCLEOTIDE POLYMORPHISMS AND MUTATIONS ON ALPHA-2-MACROGLO  
 ; FILE REFERENCE: 37481-3323  
 ; CURRENT APPLICATION NUMBER: US/10/292, 081A  
 ; PRIOR FILING DATE: 2002-11-08  
 ; PRIOR APPLICATION NUMBER: 60/337434  
 ; NUMBER OF SEQ ID NOS: 15  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 9  
 ; LENGTH: 1500  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 US-10-292-081A-9  
 Query Match 19.4%; Score 1428; DB 12; Length 1500;  
 Best Local Similarity 28.7%; Fred. No. 6.7e-108;  
 Matches 458; Conservative 270; Mismatches 567; Indels 300; Gaps 62;  
 QY 12 LVCVCTALAAVAGRPFLVTPAGIIRPGCNVTIGVLELHSCSVYVKAELLKTASNLTV 71  
 DB 15 LVLVLPDASVSGKQYVAVLPSLIHT-ETTEKGVLLSYLANETVVASLSVSGNSL 73  
 QY 72 -SVLEAEG-VFEKSGFKTLTLPSPLSNADSE---IYELRVGTQTDILFNSSTRLSSET 126  
 DB 74 FTDLAENDVLHCVAF-----AVPKSSSNEEMFLTVQVKGPTQE--FKKRTTVWKN 124  
 QY 127 KRISVFIOTDRLKYPKQEVYFRIVTLFSDRKYKTSNLIL-IDPKSNLLOMLSOQSD 185  
 DB 125 EDSLVEVQTDKSIYPGQTVKXRVVSMDEHNPINELLPLVYIDPKNRIAQWQSFQLE 184  
 QY 186 LGVISTKQLSHPFLIGDMSIOVQVND--QTYGOSFOYSEVVLPRFEVTLQTPLYCSMNS 243  
 DB 185 GGLKQSPFLSSEPPQSGYKVVQKKSGRTEH-PFTVEEFVLPRFEVQVTPKTIITLLE 243  
 QY 244 KHLNGTITAKTYTGKPVKGVDTLTFELPL-----SFMGK 276  
 DB 244 EEMNVSVGGLVTVGKFPVGHVTVSICRKYSDASDCHGSDQAFCEKFSQGLNSHGCIFYQ 303  
 QY 277 KKNITKTKINGSANFSNDEEMKAVMDSNGLSLEYLDLSFPGVEIITL-----TVTESV 331  
 DB 304 VK--TKVFOLKR-----KEYEMK-----LHTEAQIOEGVTELVLTGRQSSSEITRTI 347  
 QY 332 TGISNVSTNVFVKOHDIIEFPDYTVLKPSLNFATVAVKTVRADGNQGLTEERRNNVY 391  
 DB 348 TKLS-FVAVDSHFROG--IPFG-----QVRLVDGKVPILP---NKVI 384  
 QY 392 TVTORNTYEWSGNSGNQKMAVQ-KINTY--VPOSGTFKIEPP-----ILSDS 438  
 DB 385 FL-RGNENANY--SNATTDHGLVQFSLINTNNMGTSITVAVNVKDRSPCYGYQWVSEBH 441

```

QY 439 SELQKAYFLGSKSMVAHSLPKSPSKTYIQLKTRDENIKVS-----PFELVSGNK--R 492
D 442 EEAHTAY-----LVFSPSKSFVLEHPSHELPCGHQTQVQAHYILNGTILG 489
QY 493 LKELS---YMWVSROGLVAVG-----KONST-----MBSLTRENSMTPRACYIVYIEDDG 540
D 490 LKLSFYLLIMAKGIVTGTGHLVKGEDMKHPSISIPKSDIAPARLLIYAVLPTG 549
QY 541 EIIDVLKIPVOLFKNKIKLYMSKVKAPESEKYSLSISTYOPDSIYGVIVAVDSKNLM- 599
D 550 DVIGDSAKYDVENCILANKVDSLSPSSQSLPASHAKRYTA--POVCALRAVDSDVLML 608
QY 600 ---NAS-----NDITMENVHELELYNTGY-YLGMFMS-----FA 631
D 609 PDALSNASVYVNLPEKDLTGFPGLNQDNECINRHNVYINGITLYPVSTNEKDMYS 668
QY 632 VFQECGLVLTDAULTKYIDGVYDNE-----YAREMEENEGHIVDHPSL 680
D 669 FLEDMGLKAFNNSKIRKPCPOLQOYEMHGPGLRVGFSYDVGRGHARLVHEE--- 725
QY 681 GSPH---YRKHPETWIMDLTNMGYRYLOFEVTVPPDSISWATGVISEDGLGILT 737
D 726 ---PHTETVRKYFPEWTIMDLVYVNASAGVAVGTVPDTTEMKAGAFCLSEDLGLISS 782
QY 738 TPVELQAFQPPFIFLNFYSYIRGEEPALERTINYLKDATEVYIIEKSDKPDILMTSS 797
D 783 T-ASLRAPQPFVELTWPYSYIRGEAFILKATVNLTKPCIRVSGVQLEASPAFLAVPEK 841
QY 798 E-----INATCHOQTLVPSDEGATVLPFIPPHLGELPIVTVLMS-----PTA 841
D 842 EDAHHCICANGRO-----TVSWAVTPKSLGVNFTVSALEASOELCSTEVPSV 890
QY 842 ---SDAVTOMIIVYKAGIEKYSOSILDLTLNRLSTLKTSPSPPTVYTGSEVQ 896
D 891 PEHGRKDVIVKPLVEBGELEKETTFNSLCPSSGEVS---EELSLKLPVNVESARAS 947
QY 897 ITAIGDVLGSPSINGLASLIRMPYCGGEQNMINFAPNIYILDYLTKKQLTDLNLEKALSF 956
D 948 VSVLGDILGSAWQNTQNLQWPGYCGGEQNVLPAPNIYVLDYLNQOQLTPEIKSKAIGY 1007
QY 957 MRQGYRELLYQREDEGSFASG-NYDS--GSTWLSAPLRCFLLEADPIIDQVNLHRTY 1014
D 1008 LNTYQOROLNYKHVDSYSTGEERYGRQGTWLTAFILKFAQARAYIFIDEAHITQAL 1067
QY 1015 TWLKGHOSNGEFPDQGVHSELOGNKSVPVLTAVYVTSLSGVRKQOPIDVQESHF 1074
D 1068 TWSORQKONGCFSSGSLNNNAIKGVEDEVITLSAYITTALEIPLTVHPPVKNALPFC 1127
QY 1075 LES-----EFSRGISDNYTLALITYALSVS--PKAKEALNMLTWAEQSGMOW--- 1124
D 1128 LESAMKTAQEGDHG-SHYTTKALLAYAFALAGNDKREVVLKSLNEBAVKKONSVMHERP 1186
QY 1125 VSSSEKSDSKQPR--SLDIEVAAYALISHLQFQ---TSEGIP---INMWSLRQNSL 1175
D 1187 QKPRAPVGHFEYPOAPSAVEVEMTSYVLA--YLTAQAPVTSBDLTSAITVIMIKQONAQ 1245
QY 1176 GGFASDTDTYALAKALSEFALNMTERTNIQVYVYTGSSPEPLAVQ----- 1222
D 1246 GGFSTQDTYVALHALSKYGAATFT-RTGKAQAVYILOSSTGFSKRFVDNNRLLLOQVS 1304
QY 1223 ---PMAVNISANGFPAICQNLVYVNVKASGSSRRRSRISNOEAFDLDVAVK---ENKD 1275
D 1305 LPELPGESYMKVTEGCVYLOTSLKYNL-----LPEKEFPFALGVQLPQTC 1353
QY 1276 DLN-HVDLNVCTSPSGPG--RSGMALMEVNLISGFVWPSAISL---SETVKVVEYDHG 1328
D 1354 EPKAKHTFQIQLSYVYSGSRASNMAYIDVVMVSGFIPLKFTVAKLERSNHNSTEVSN 1413
QY 1329 KLANIYLSVNETQCVNIPARANKVSVTODASVYVYVPRQAVRSNYSVEKLSSCD 1388
D 1414 HLVIILDKVSNQTLSPFTVYLDVVPVRLKPAIVKVDYB-----TGDQLQLSTLM 1464

```

```

QY 1389 LCS-----DVQCRPCEDGAGSGHHHSVITF 1416
D 1465 LLAAKILEMLEDHKAEXCFAGVLFSELRHRRHPLY 1499

RESULT 15
US-09-981-151A-10
; Sequence 10, Application US/09981151A
; Publication No. US20030212256A1
; GENERAL INFORMATION:
; APPLICANT: Edinger, Shlomit R
; APPLICANT: Gerlach, Valerie
; APPLICANT: MacDougall, John R
; APPLICANT: Malyanekar, Muriel M
; APPLICANT: Smithson, Glenda
; APPLICANT: Miller, Isabelle
; APPLICANT: Peyman, John A
; APPLICANT: Stone, David J
; APPLICANT: Gunther, Erik
; APPLICANT: Ellerman, Karen
; APPLICANT: Shinkets, Richard A
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Guo, Xiaojia
; APPLICANT: Paturajan, Meera
; APPLICANT: Taupier Jr, Raymond J
; APPLICANT: Burgess, Catherine E
; APPLICANT: Zerhusen, Bryan D
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Spletz, Kimberly A
; APPLICANT: Gangoli, Baha A
; APPLICANT: Fernandes, Elma R
; APPLICANT: Gorman, Linda
; TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-168
; CURRENT APPLICATION NUMBER: US/09/981,151A
; CURRENT FILING DATE: 2001-10-16
; PRIOR APPLICATION NUMBER: 60/241,040
; PRIOR FILING DATE: 2000-10-17
; PRIOR APPLICATION NUMBER: 60/241,058
; PRIOR FILING DATE: 2000-10-17
; PRIOR APPLICATION NUMBER: 60/241,063
; PRIOR FILING DATE: 2000-10-17
; PRIOR APPLICATION NUMBER: 60/241,243
; PRIOR FILING DATE: 2000-10-17
; PRIOR APPLICATION NUMBER: 60/242,152
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/242,482
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/242,611
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/242,612
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/242,880
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: 60/242,881
; PRIOR FILING DATE: 2000-10-24
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 160
; SOFTWARE: Patentln Ver. 2.1
; SEQ ID NO 10
; LENGTH: 1492
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-981-151A-10

Query Match 19.3%; Score 1421.5; DB 12; Length 1492;
Best Local Similarity 27.2%; Pred. No. 2.3e-107;
Matches 419; Conservative 269; Mismatches 559; Indels 293; Gaps 44;

```

```

QY 20 LAYAPG-----PFLYTABRIR-PCGNVYIGVLEHCSQYTVKAEILKTSNLTIVSV 73
D 10 LALSFAIELPMLVYVTLTPARLNF-----SVQKCVLDLSPGYSVDKFTVLTETKDKTKQL 65

```



```

QY 74 LEAEVFEKSGFKTLTLPBPLPLNSADEIYELKVTGRTQDELLFNSNTRLSFETKRI SVPI 133
Db 66 LEVSGKHLKHLHCISLFPVPAGGTEEVKATIRVSG-VGNNISFEEKCKTLIRQGGTFV 124
QY 134 QTDKALYKROEKYKPIVTLFSDFKPKYSINIL-1KDPKSNLIOQMLSQSDIGVIST 192
Db 125 QTDKPLVYTGQOYFRIVTMDSNFVVDNKTSMVELQDPNSNRILQMLVEVEQGVLDIS 184
QY 193 POLSSHPILGWSIOQVNDQTYOSFOVSEY-----VLPKFEVTLQTPLYCSM 241
Db 185 POLAPEAMLGYT--VAVAEKGTFGFSVEEYVLSFPLLLSSVLPKFEVEVEPELST 242
QY 242 NSKHLNGITTAKTGYKPKVGDVTLFPLSLFSGKKKNI TKTPKINGANSFNDDEKMN 301
Db 243 VQESPLVKICCRYYTGKPMIGAVOVSV--QKAN-----TYWREVEREQ 285
QY 302 VWDSSNGLSEYLD--LSSP-----GPVELLTYYTESVTGIRSNSTNIF 343
Db 286 LPDKCRNLISGQTDKTCGFSAPVDMATEDLIGYAVSHQINI VAVVEEGVEBANNTQNIY 345
QY 344 FKOHDIIEFPDYTVLVKPSLNFATVKTBRADGNOLTEERRNNVI-----TV 393
Db 346 ISPOGSMWFEDTSTNYHNFPFSGKMLKFPQGVLPCKNHLVFLVITGTNGTNGTIV 405
QY 394 TORN-----YTERWSSNSNGQKAEAVOKINTVPOGTFKIE-----FPLBDS 438
Db 406 TUNNGIAPFTELETSGWNGTD-----VSLGKFKOMEDLVVNPQVPRYYON 450
QY 439 SELQKAVYLGSKSSNAVSLFKSPKTYIQLKTRDENIKVSPFELV-----SG 489
Db 451 AYLHLRPFYSTRSFLGIHRL-----NGPLKCGQPOEVLVDYIIDPADASP 496
QY 490 NKLKELSTYVVRSGOLVAVGKON-----STMFLTPENSGWTPKACIVYYIEDD 539
Db 497 DDEI-SFSTYLLGKSGLVMEQKHLNKKKGLKASFSLSLTFSRLAPPSLVITAIIPS 555
QY 540 GEIISDVLEKIPVQLVFNKIKLYWSKVKAPESEKVSIRISVTPDSIVGIVAVDKSNLM 599
Db 556 GGVVADKIQFSVEMCFDN-----QQLPGAVEVLQLOAA-PSGLCALRAVDESVLIL 605
QY 600 NASNDITMENNVHELELYNTGYLGMF----- 626
Db 606 RP-----DRELSNRSVY-GMFPWGHYPYQVAEYDQCPVSGPMPDPLIDPM 653
QY 627 -----MNSPAVROEGGLWLTDTANLTQYIDGVNDAEYA----- 661
Db 654 POGHSSORSITWRPSESEGTDLFSFRVGLKILSNKIKKP-VDCSHRSPBYSTAMGG 712
QY 662 ---ERFMENEGHIVDIDHFSLGSSPHVKHPPETWIMLDTMNGYRIYOEFEVTVPSIT 718
Db 713 GHPEAFESSTPLHQAE-----DSQVQYFPEFTWIMDLFPI GNSGKEAVHVTVPDALT 764
QY 719 SWVAAGFVISEDLGGLTTTPEVLOAFQFFIFLNPYSVIRGEBFALEITTFNYIKDAT 778
Db 765 EKKAMSFCTSOGRGFLSPT-VGLTAFKPPFDLTLPSVVRGSESEFLTATIFNYIKDCI 823
QY 779 EYKVIIEKSDKFDILMTSEINATGHOQLVPSDEGATVLPPIRPTHGELPIITYAL- 837
Db 824 RYQTDLANSHYQLESWADS-----QISSCLCADDAKTHWNITAVKLGHNFTISTYKI 877
QY 838 -----SPTASDAVTOMILVKAIGIEKYSQSILDLTDNRLOSTLKTLSF 882
Db 878 LUSNEPCGQKGFVPQKGRSDTLIKPVLYKPEGVLEKTHSSLLCPKGGKVAS--ESVSL 935
QY 883 SPPPVTVTSGSERVOITAGDVLGFSINGLASLIRMPYGGCEQNMIFADNIYILDYTK 942
Db 936 ELPPVDIVPSTKAYVTVLGDIMGTAQLONDGLVOMPSGCEQNMVIFABIIYVLYLEKA 995
QY 943 KQITDLKXKALSPMGOYORELLYORHDSFSAFENVDPSCSTWLSAFVLCFLEADPY 1002
Db 996 GLITEIRIRAVGFLGIGQKELMYHSGSYSAFERDNGNNTWLTAFVTKCFQOQKCF 1055
QY 1003 IDIDQNVHLRTYTWLKGHQKNGEFPWDPGRVHSELQGNKSPVLTAVITVSLGYRKY 1062

```

```

Db 1056 IFIDPKXNIDALKWAGQOLPSGCYANVGNLHTAMKGVVDDEVSLTAVYTPALLMKGCD 1115
QY 1063 QPNIDVOSIHFLESEFSGISDNVTLTALITYALSSVSGSPRAKE-ALNMLTWRAQEGCM 1121
Db 1116 VDDPVSOGKCLKNS-RTSTNLYTQNLHLIYISLAGEMDIRNILLQLOQALISGES 1174
QY 1122 QFWSSSEKSLDS--W-OPRSIDIEVAAYALLSH-----PLOFQTSSEGIPIRMWLSRORN 1173
Db 1175 IYMSQKPTPSSNASPMSPEPAVDVELTAVALLAQUTKSLQKELAKKTSIYAMLAKQHN 1234
QY 1174 SLGFSASTODTTVALKALSEPA--ALNTERNTIOVTVT-----GPSSPEPLAVQ-- 1222
Db 1235 AYGGFSTODTVVALQALAKYATTAAMPSEINLVKSTENFORTFNIOQVNRLVFOODT 1294
QY 1223 ---PMANISANGGFACQOLNVVYVVKASGSSRRRRSIONQEAFLDVAVK---ENK 1274
Db 1295 LPNVGMYTLESGGQCVYVQTVLRKYNILPP-----TNMKTFSLSVEIGKARCQOP 1345
QY 1275 DDLNHDVNLNCTSFSGP-GRSGMALMEVNLISGF--WVPSBAISSE-TYKKVEYDHQKL 1330
Db 1346 TSPRSLTLTHTSYGSRSSSNMAIVEYKMLSGFSPMEGCTQOLLOPVLKKEVERGTOTL 1405
QY 1331 NLYUDSVNETQPCVNI PAVRNFKXSNQDASVSIYDYEYEP 1370
Db 1406 NIYDELIKNTQTYFTISQSVLVTNLKPKATIKYDYDYL 1445

```

Search completed: January 15, 2004, 18:19:33  
Job time : 51 secs

**THIS PAGE BLANK (uspto)**

```

      RESULT 1
      US-09-016-434-1174
      Sequence 1174, Application US/09016434
      Patent No. 6500938
      GENERAL INFORMATION:
      APPLICANT: Janice Au-Young
      APPLICANT: Jeffrey J. Sellhammer
      TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF SIGNALING
      TITLE OF INVENTION: PATHWAY GENE EXPRESSION
      NUMBER OF SEQUENCES: 1430
      CORRESPONDENCE ADDRESS:
      ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
      STREET: 3174 PORTER DRIVE
      CITY: PALO ALTO
      STATE: CALIFORNIA
      COUNTRY: USA
      ZIP: 94304
      COMPUTER READABLE FORM:
      MEDIUM TYPE: Floppy disk
      COMPUTER: IBM PC compatible
      OPERATING SYSTEM: PC-DOS/MS-DOS
      SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
      CURRENT APPLICATION DATA:
      APPLICATION NUMBER: US/09/016,434
      FILING DATE: HEREWITH
      CLASSIFICATION:
      PRIOR APPLICATION DATA:
      APPLICATION NUMBER:
      FILING DATE:
      CLASSIFICATION:
      ATTORNEY/AGENT INFORMATION:
      NAME: Zeller, Karen J.
      REGISTRATION NUMBER: 37,071
      REFERENCE/DOCKET NUMBER: PA-0002 US
      TELECOMMUNICATION INFORMATION:
      TELEPHONE: (650) 855-0555
      TELEFAX: (650) 845-4166
      INFORMATION FOR SEQ ID NO: 1174:
      SEQUENCE CHARACTERISTICS:
      LENGTH: 4079 base pairs
      TYPE: nucleic acid
      STRANDEDNESS: single
      TOPOLOGY: linear
      IMMEDIATE SOURCE:
      LIBRARY: GENBANK
      CLONE: g177869
      US-09-016-434-1174
  
```

[illegible]

```

: NAME/KEY: mat peptide
: LOCATION: (113)..(4468)
US-09-241-606-1

Query Match      2.6%; Score 110; DB 4; Length 4577;
Best Local Similarity 50.2%; Pred. No. 3.3e-21;
Matches 303; Conservative 0; Mismatches 295; Indels 6; Gaps 1;

QY 2649 ATTTCCTCTTAATACAGTACTGGCAGTGAAGAGTTCAATCACTCCATTTGGAGATGT 2708
DB 2845 ACTCCCAACAAATGTGTAGAAAGATGTGCCGCCGAGCTTGCTGCTCACTTTTGGAGACAT 2904
QY 2709 TCTTGATCTTCCATCAATGGCTTACGCTCAATGATTCGGATGGCTTATGGCTGTGTA 2768
DB 2905 ATTAGGCTTGCCATGCAAAACACAAAAATCTTCTCAAGTGGCTTAATGCTGTGAGA 2964
QY 2769 ACAGACATGATAAATTTTGCTCCAAATATTTCATTTTGGATTATGTGACATAAAAGAA 2828
DB 2965 GCAGAAATATGGTCTCTTGTGCTCTTAACATCTAATGTAAGTGAATATCTAAATGAACACA 3024
QY 2829 ACAACTGACAGATATTTGAAAGAAAAGCTCTTCAATTATGAGGCAAGTTTCCAG 2888
DB 3025 GCACCTTAATCTCAAGATCAAGTCCAAAGGCCATGGCTATCTCAACACTGGTAAACAGAG 3084
QY 2889 AGAAGCTCTCATAGAGGGGAAGATGGCTCTTTCAGTGGCTTTGGG-----AATTATGA 2942
DB 3085 ACACTGTAATCAAAACACTATGATGGCTCTCAACAGACCTTTGGGAGCAGATATGGCAG 3144
QY 2943 CCTTCTGGAGCACTTGTTGTTCAGCTTTTGTTTAAGATGTTTCTTGAAGCCGATCC 3002
DB 3145 GAACCAAGGGCAACCTGGGCTCAAGGCTTTGTTCTGAAGACTTTTGGCCCAAGCTCGAGC 3204
QY 3003 TTACATGATATTTGATCAGAATGTGTTACACAGAACATACACTTGGCTTAAAGACATCA 3062
DB 3205 CTACATCTTTCATCATGATAGAGACACATTAACCAAGCCCTCATATAGGCTTCTCCAGAGCA 3264
QY 3063 GAATCCAGCGGTGAATTTTGGATCCAGGAAGATTCATATGATGAGCTTCAAGTGG 3122
DB 3265 GAAAGACATATGGCTGTTTCAGAGAGCTTGGGTCACTCTCAACATATGCCATTAAGGAGG 3324
QY 3123 CATAAAAAGTCCAGTAACACTTACAGCCTAATTTGTAATCTTCTCTGGGATATAGAA 3182
DB 3325 AGTGAAGATGAATGACCCCTCTCGGCTAATACCATGGCCCTTCTGGAGATTCCTCT 3384
QY 3183 GTATCAGCTAATCATTTGATGTGCAAGAGTCTATCCATTTTGGAGTCTGAATTCAGTAG 3242
DB 3385 CACAGTCACTCAACCTGTGTGCGCAATAGCCCTGTTTGCTGGAGTCAAGCTTGAAGAC 3444
QY 3243 AGGA 3246
DB 3445 AGCA 3448

RESULT 3
US-08-232-463-14/C
: Sequence 14, Application US/08232463
: Patent No. 5670367
: GENERAL INFORMATION:
: APPLICANT: DORNER, F.
: APPLICANT: SCHREIFLINGER, F.
: APPLICANT: FALKNER, F.G.
: TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
: NUMBER OF SEQUENCES: 52
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Foley & Lardner
: STREET: 1800 Diagonal Road, Suite 500
: CITY: Alexandria
: STATE: VA
: COUNTRY: USA
: ZIP: 22313-0299
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
/

```

```

? OPERATING SYSTEM: PC-DOS/MS-DOS
? SOFTWARE: Patent In Release #1.0, Version #1.25
? CURRENT APPLICATION DATA:
?   APPLICATION NUMBER: US/08/232,463
?   FILING DATE:
?   CLASSIFICATION: 435
? PRIOR APPLICATION DATA:
?   APPLICATION NUMBER: US/07/935,313
?   FILING DATE:
?   APPLICATION NUMBER: EP 91 114 300,6
?   FILING DATE: 26-AUG-1991
? ATTORNEY/AGENT INFORMATION:
?   NAME: BENT, Stephen A.
?   REGISTRATION NUMBER: 29,768
?   REFERENCE/DOCKET NUMBER: 30472/114 IMMU
? TELECOMMUNICATION INFORMATION:
?   TELEPHONE: (703)836-9300
?   TELEFAX: (703)683-4109
?   TELEX: 899149
? INFORMATION FOR SEQ ID NO: 14:
? SEQUENCE CHARACTERISTICS:
?   LENGTH: 7218 base pairs
?   TYPE: nucleic acid
?   STRANDEDNESS: single
?   TOPOLOGY: linear
? IMMEDIATE SOURCE:
? CLONE: PTZgpt-F15
US-08-232-463-14

```

MOLECULE TYPE: DNA (genomic)  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 9..4961  
US-08-447-411-1

Query Match 1.1%; Score 46.4; DB 1; Length 5211;  
Best Local Similarity 50.4%; Pred. No. 0.012;  
Matches 113; Conservative 0; Mismatches 111; Indels 0; Gaps 0;

QY 3376 TCATCAGATCCCAATTTCTGACTCTGAGCAGCAGCCTCCCTGATATGAAAGTTGCA 3435  
DB 3609 TCAACAGAGAGAGATCGTTGGAGAAATATATATCTGCAACCCATATATGAGGCACT 3668  
QY 3436 GCCTATGACATGCTCTGACACTTCTCAATTTGAGACTTGAGGGAAATCCCAATTATG 3495  
DB 3669 TCCATGCTCTTTGGGCTCTGCTGAAATGAAATTTGCTGAGGTGGGCTCTGTGTC 3728  
QY 3496 AGGTGGCTAAGCAGCAAGAAATAGCTTGGGTGTTTGCATCTACTCAGATACACT 3555  
DB 3729 AGATGGCTGATAGATCAGAAATATATGAGGGAACATATGAGCAAAACCAAGCAAGTT 3788  
QY 3556 GTGCTTTAAAGCTCTGTCTGATTTGACCCCTAATGATAC 3599  
DB 3789 ATGGTGTTCAGAGCTTGTGTAATATGAGATTCAGATGCTTAC 3832

RESULT 6  
US-09-241-606-5  
Sequence 5, Application US/09241606  
Patent No. 6472140  
GENERAL INFORMATION:  
APPLICANT: Tanzi, Rudolph E.  
APPLICANT: Kovacs, Dora  
APPLICANT: Saunders, Aleister J.  
TITLE OF INVENTION: Alpha-2-Macroglobulin Therapies and Drug Screening Methods for  
FILE REFERENCE: 0609.446003  
CURRENT APPLICATION NUMBER: US/09/241.606  
NUMBER OF SEQ ID NOS: 27  
SOFTWARE: Patentin Ver. 2.0  
SEQ ID NO 5  
LENGTH: 333  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (1)..(333)  
OTHER INFORMATION: A Binding Domain  
US-09-241-606-5

Query Match 1.0%; Score 44; DB 4; Length 333;  
Best Local Similarity 63.0%; Pred. No. 0.0092;  
Matches 68; Conservative 0; Mismatches 40; Indels 0; Gaps 0;

QY 3481 GGAATCCCAATTAAGGTGCTTAAGCAGCAAGAAATACCTGGGTGTTGCATCT 3540  
DB 19 GCAACCAACATCGTGAAGTGCATCAGAACGAGCAAGAAATGCCAGGGGGTTCTCTCC 78  
QY 3541 ACTCAGATACCACTGCTTTAAAGCTCTGTCTGATTTGAGCC 3588  
DB 79 ACCAGAGACACAGTGTGCTCTCCATGCTCTCTCAAAATGAGGCC 126

RESULT 7  
US-09-241-606-3  
Sequence 3, Application US/09241606  
Patent No. 6472140  
GENERAL INFORMATION:  
APPLICANT: Tanzi, Rudolph E.  
APPLICANT: Kovacs, Dora  
APPLICANT: Saunders, Aleister J.

TITLE OF INVENTION: Alpha-2-Macroglobulin Therapies and Drug Screening Methods for  
FILE REFERENCE: 0609.446003  
CURRENT APPLICATION NUMBER: US/09/241.606  
CURRENT FILING DATE: 1999-02-02  
NUMBER OF SEQ ID NOS: 27  
SOFTWARE: Patentin Ver. 2.0  
SEQ ID NO 3  
LENGTH: 750  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (1)..(750)  
OTHER INFORMATION: A /LRP Binding Domain  
US-09-241-606-3

Query Match 1.0%; Score 44; DB 4; Length 750;  
Best Local Similarity 63.0%; Pred. No. 0.016;  
Matches 68; Conservative 0; Mismatches 40; Indels 0; Gaps 0;

QY 3481 GGAATCCCAATTAAGGTGCTTAAGCAGCAAGAAATACCTGGGTGTTGCATCT 3540  
DB 19 GCAACCAACATCGTGAAGTGCATCAGAACGAGCAAGAAATGCCAGGGGGTTCTCTCC 78  
QY 3541 ACTCAGATACCACTGCTTTAAAGCTCTGTCTGATTTGAGCC 3588  
DB 79 ACCAGAGACACAGTGTGCTCTCCATGCTCTCTCAAAATGAGGCC 126

RESULT 8  
US-09-356-952-11  
Sequence 11, Application US/09356952  
Patent No. 6117663  
GENERAL INFORMATION:  
APPLICANT: Borjesson, Ann  
APPLICANT: Margalit, S. M.  
APPLICANT: Bor-Sogli, Dafna  
APPLICANT: Cole, Philip  
APPLICANT: Kurijyan, John  
TITLE OF INVENTION: A CRYSTAL OF A RAS-SOS COMPLEX AND METHODS OF USE  
FILE REFERENCE: 600-1-228N  
CURRENT APPLICATION NUMBER: US/09/356.952  
CURRENT FILING DATE: 1999-07-19  
EARLIER APPLICATION NUMBER: 60/093.631  
EARLIER FILING DATE: 1998-07-21  
NUMBER OF SEQ ID NOS: 14  
SOFTWARE: Patentin Ver. 2.0  
SEQ ID NO 11  
LENGTH: 5398  
TYPE: DNA  
ORGANISM: Saccharomyces cerevisiae  
US-09-356-952-11

Query Match 1.0%; Score 41.4; DB 3; Length 5398;  
Best Local Similarity 47.2%; Pred. No. 0.37;  
Matches 126; Conservative 0; Mismatches 141; Indels 0; Gaps 0;

QY 802 TTTTACCTTATCCCTTTGGGAAAGAGAAATATTAACAAAATTAAAGATAAT 861  
DB 2671 TTTTACCTTATCCCTTTGGGAAAGAGAAATATTAACAAAATTAAAGATAAT 2730  
QY 862 GGAATCCCAATTAAGGTGCTTTAAAGCTCTGTCTGATTTGAGCC 921  
DB 2731 AGAAGAGAGAGAGATCGTGAAGTGCATCAGAACGAGCAAGAAATGCCAGGGGGTTCTCTCC 2790  
QY 922 GGAATCCCAATTAAGGTGCTTTAAAGCTCTGTCTGATTTGAGCC 981  
DB 2791 CAAATTTTAAAGATTAATATGCTACAGGGAACATTTAAATTAAGTAAACC 2850  
QY 982 ACAGATCAGTACAGTATTTCAAGAAATGTAAGCACTAATGTGTTCTTCAAGCAAT 1041



Db 2851 AAAAGCAGATTGAGATTGGAAATTAATTCAGACATACGACAAATTAATCAGAAAT 2910  
Oy 1042 GATTACATCATTTGAGTTTGGATTAT 1068  
Db 2911 GTTTACTATTGAGATTACTGAGAAAT 2937

RESULT 9  
US-08-916-421B-1/c  
Sequence 1, Application US/08916421B  
Patent No. 6503729  
GENERAL INFORMATION:  
APPLICANT: Bult et al.  
TITLE OF INVENTION: Complete Genome Sequence of the Methanogenic Archaeon, Methanoco  
Patent No. 6503729  
FILE REFERENCE: PB275  
CURRENT APPLICATION NUMBER: US/08/916,421B  
CURRENT FILING DATE: 1997-08-22  
PRIOR APPLICATION NUMBER: US 60/024,428  
NUMBER OF SEQ ID NOS: 3  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 1  
LENGTH: 1664976  
TYPE: DNA  
ORGANISM: Methanococcus jannaschii  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (28222)..(28222)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc\_feature  
LOCATION: (28257)..(28258)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc\_feature  
LOCATION: (84773)..(84773)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc\_feature  
LOCATION: (84808)..(84808)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc\_feature  
LOCATION: (84812)..(84812)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc\_feature  
LOCATION: (98150)..(98150)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc\_feature  
LOCATION: (98159)..(98159)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc\_feature  
LOCATION: (98239)..(98239)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc\_feature  
LOCATION: (98266)..(98266)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc\_feature  
LOCATION: (98343)..(98343)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc\_feature  
LOCATION: (10398)..(10398)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc\_feature  
LOCATION: (148548)..(148548)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc\_feature  
LOCATION: (16385)..(16385)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc\_feature  
LOCATION: (19189)..(19189)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc\_feature  
LOCATION: (19195)..(19195)  
OTHER INFORMATION: n equals a, t, c, or g

NAME/KEY: misc\_feature  
LOCATION: (231980)..(231980)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc\_feature  
LOCATION: (234187)..(234187)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc\_feature  
LOCATION: (234220)..(234220)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc\_feature  
LOCATION: (234814)..(234814)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc\_feature  
LOCATION: (309398)..(309398)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc\_feature  
LOCATION: (309418)..(309418)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc\_feature  
LOCATION: (312837)..(312837)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc\_feature  
LOCATION: (312993)..(312993)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc\_feature  
LOCATION: (319226)..(319226)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc\_feature  
LOCATION: (559167)..(559167)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc\_feature  
LOCATION: (559241)..(559241)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc\_feature  
LOCATION: (600992)..(600992)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc\_feature  
LOCATION: (622708)..(622708)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc\_feature  
LOCATION: (657081)..(657081)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc\_feature  
LOCATION: (657203)..(657203)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc\_feature  
LOCATION: (674435)..(674435)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc\_feature  
LOCATION: (682442)..(682442)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc\_feature  
LOCATION: (713652)..(713652)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc\_feature  
LOCATION: (741684)..(741684)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc\_feature  
LOCATION: (779455)..(779455)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc\_feature  
LOCATION: (779676)..(779676)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc\_feature  
LOCATION: (855539)..(855539)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc\_feature  
LOCATION: (871619)..(871619)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc\_feature  
LOCATION: (1084830)..(1084830)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc\_feature

```

Query Match      1.0%; Score 41.4; DB 4; Length 1664976;
Best Local Similarity 52.0%; Pred. No. 19;
Matches 93; Conservative 0; Mismatches 86; Indels 0; Gaps 0;

Qy      3725 TTGTATATATAATGTGAAGCGCTTCTGGCTCTTCTAGAGAAGAGATCTATCCAAATCAAG 3784
Db      539968 TTGGAGCTACTGATATAGCGCTTACATCTATGCAACAGAGAAATCATGATTAAAGTCCAA 539905

Qy      3785 AAGCCTTGATTTTATGATGTTCGTCTTAAAGAAATTTAAGATGATCTCAATCATGTGCATT 3844
Db      539908 AAACAATTTAGGTAGATATATGTTGGAAAAAATGAAAAATGTTCTCCCAAGAATATTTGTT 539844

Qy      3845 TGATGTGTGTACAACTTTTCGGGCCCGGTAGAGATGGCATGCGCTTATGGAAGTT 3903
Db      539848 TAAGCGTTTGTAAAGAAATTTGGAGAGAGAGCAACAATCATGTGCTATTGATATGCT 539790

RESULT 10
US-09-601-198-80
; Sequence 80, Application US/09601198
; Patent No. 6531583
; GENERAL INFORMATION:
; APPLICANT: Cassell, Gail H.
; APPLICANT: Chen, Ellison Y.
; APPLICANT: Glass, Jennifer S.
; APPLICANT: Glass, John I.
; APPLICANT: Heiner, Cheryl R.
; APPLICANT: Lefkowitz, Elliot
; TITLE OF INVENTION: NUCLEIC ACID PROBES AND METHOD FOR DETECTING UREAPLASMA
; FILE REFERENCE: UAB-13452/22
; CURRENT APPLICATION NUMBER: US/09/601,198
; CURRENT FILING DATE: 2000-12-08

```

Query Match 0.9%; Score 40; DB 4; Length 1581;  
Best Local Similarity 43.8%; Pred. No. 0.4;  
Matches 175; Conservative 0; Mismatches 225; Indels 0; Gaps 0;

## RESULT 11

```

US-09-601-198-60/c
; Sequence 60, Application US/09601198
; Patent No. 6531583
; GENERAL INFORMATION:
; APPLICANT: Cassell, Gail H.
; APPLICANT: Chen, Ellison Y.
; APPLICANT: Glass, Jennifer S.
; APPLICANT: Glass, John I.
; APPLICANT: Heiner, Cheryl R.
; APPLICANT: Lefkowitz, Elliot
; TITLE OF INVENTION: NUCLEIC ACID PROBES AND METHOD FOR DETECTING UREAPLASMA
; TITLE OF INVENTION: UREALYTICUM
; FILE REFERENCE: UAB-13452/22
; CURRENT APPLICATION NUMBER: US//09/601,198
; CURRENT FILING DATE: 2000-12-08
; PRIOR APPLICATION NUMBER: 60/073,189
; PRIOR FILING DATE: 1998-01-30
; NUMBER OF SEQ ID NOS: 181
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 60
; LENGTH: 15016
; TYPE: DNA
; ORGANISM: Ureaplasma urealyticum
US-09-601-198-60

```

Query Match	0.9%	Score 39.4;	DB 4;	Length 15016;
Best Local Similarity	48.1%	Pred. No. 2.9;		
Matches 112; Conservative	0;	Mismatches 121;	Indels 0;	Gaps 0;

QY 703 ACACCTATATGTTCTATGATTCGATTAAGCTTAATGTCACATCAGCGCAAGTAT 762  
DB 2859 ACACCTATATGATTAAGATATTAATTAAGATATTCACGCAAAATATTCAGTATTT 2800  
QY 763 ACATATGGGAGCCAGTGAAGGAGCGTAACGCTTACCTTTTACCTTATCCCTTTGG 822  
DB 2799 GAACTATACATACATTAATTAATGAAATTTTAATAGCAATGCTTTTATCTTTAAT 2740  
QY 823 GGAAGGAGGAAATATTTACAAATTCATTTAGATTAATGATGCAAACTTCTTTT 882  
DB 2739 AATTAGAGCCCTAATATTTAATATAAATTAATTAATTAATGATTTGAGCAAAACCACT 2680  
QY 883 AATGATGAAGATGAATAATGTAATGATTCCTTAATGATGACTTTTGATATA 935  
DB 2679 AAGGCTATGCTAATTAATTAATCAAAAGCTAATAAATGATTAATTAAGAAA 2627

## RESULT 12

US-09-350-756-9  
; Sequence 9, Application US/09350756  
; Patent No. 6495143  
; GENERAL INFORMATION:  
; APPLICANT: U.S. Army Medical Research Institute for Infectious Diseases  
; APPLICANT: John S. Lee  
; APPLICANT: Peter Pushko  
; APPLICANT: Michael D. Parker  
; APPLICANT: Jonathan F. Smith  
; APPLICANT: Mark T. Dertzbaugh  
; TITLE OF INVENTION: Botulinum Neurotoxin Vaccine  
; FILE REFERENCE: 003/124/SAP R11D 98-21  
; CURRENT APPLICATION NUMBER: US/09/350,756  
; EARLIER FILING DATE: 1999-07-09  
; EARLIER APPLICATION NUMBER: US 60/092,416  
; NUMBER OF SEQ ID NOS: 11  
; SOFTWARE: Apple Macintosh Microsoft Word 6.0  
; SEQ ID NO 9  
; LENGTH: 1327  
; TYPE: DNA  
; ORGANISM: Clostridium botulinum  
; FEATURE:  
US-09-350-756-9

Query Match 0.9%; Score 38.8; DB 4; Length 1327;  
Best Local Similarity 49.5%; Pred. No. 0.79;  
Matches 100; Conservative 0; Mismatches 102; Indels 0; Gaps 0;

QY 1472 AACGATTGAAGGAGTTAAGCTATATGATGATCCAGGGGACAGTTGCTGCTAGGAA 1531  
DB 801 AATATATGATAGTATTTAGAGTATATGATCTTTAAAGGCGCTAAGAGCTAATGAC 860  
QY 1532 AACAAATTCACAAATGTTCTCTTTAACACAGAAATTCCTGACCTCCAAAGCCTGTG 1591  
DB 861 TACAATATGCTACAAACATTTATTTAAATTCAGTTTGTATGGGGGCAAAATTTATTA 920  
QY 1592 TAAATGCTATATATGTAAGATGATGGGAAATTAATGATGTTCTAAAAATTCCTG 1651  
DB 921 TAAAAAATATGCTCTCTGGAATAAAGATTAATTTGTTAAGAAATATGATGCTGATATA 980  
QY 1652 TTCACCTGTTTAAATAA 1673  
DB 981 TTAATGCTAGTACTTAAAAATA 1002

## RESULT 13

US-08-955-138-2/c  
; Sequence 2, Application US/08955138A  
; Patent No. 5977435  
; GENERAL INFORMATION:  
; APPLICANT: Lefebvre, Daniel D.  
; APPLICANT: Gelliaty, Kevin S.  
; TITLE OF INVENTION: PLANT PHOSPHATASES

; FILE REFERENCE: PPL97-01  
; CURRENT APPLICATION NUMBER: US/08/955,138A  
; CURRENT FILING DATE: 1997-10-21  
; NUMBER OF SEQ ID NOS: 119  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 2  
; LENGTH: 3981  
; TYPE: DNA  
; ORGANISM: SOLANUM TUBEROSUM  
US-08-955-138-2

Query Match 0.9%; Score 38.8; DB 2; Length 3981;  
Best Local Similarity 50.0%; Pred. No. 1.7;  
Matches 97; Conservative 0; Mismatches 97; Indels 0; Gaps 0;

QY 1503 ATCCAGGGGAGAGTGTGGGCTGTAGGAAACAAATTCACATAGTCTCTTTACACC 1562  
DB 3066 ATCTAGACACAAATGTTCTTAAATTTGACACCTTAATTAATTTATCTCAGAGTTAACTT 3007  
QY 1563 AGAAATTCCTTGACTCCAAAGCCTGTGTAATTTGATTAATTAATGATGAGTGGGA 1622  
DB 3006 AGAATACCTTAACCGCCACTACTTAAGTATATTTGTTATATGAACTAGCAAGTA 2947  
QY 1623 AATTATTAAGTATGTTCTAAATTCCTGTTCACTGTTTAAATAATGATTAAGT 1682  
DB 2946 AATTATTAATTTGTTTAAACCTTTTGTGAATCATTAATTAATTAAGTAAATGTT 2887  
QY 1683 AATTGAGTAAAG 1696  
DB 2886 CAATGTTGTGAAG 2873

## RESULT 14

US-09-071-035-83  
; Sequence 83, Application US/09071035  
; Patent No. 6448043  
; GENERAL INFORMATION:  
; APPLICANT: G.H. Choi  
; TITLE OF INVENTION: Enterococcus faecalis Polynucleotides and Polypeptides  
; NUMBER OF SEQUENCES: 496  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Human Genome Sciences, Inc.  
; STREET: 9410 Key West Avenue  
; CITY: Rockville  
; STATE: Maryland  
; COUNTRY: USA  
; ZIP: 20850  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage  
; COMPUTER: HP Vectra 486/33  
; OPERATING SYSTEM: MSDOS version 6.2  
; SOFTWARE: ASCII Text  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/071,035  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: A. Anders Brooks  
; REGISTRATION NUMBER: 36,373  
; TELEPHONE: (301) 309-8504  
; TELEFAX: (301) 309-8512  
; INFORMATION FOR SEQ ID NO: 83:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1579 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: linear  
US-09-071-035-83

Query Match 0.9%; Score 38.6; DB 4; Length 1579;  
Best Local Similarity 49.3%; Pred. No. 1;  
Matches 101; Conservative 0; Mismatches 104; Indels 0; Gaps 0;  
QY 1501 GTATCCAGGAGACAGTGGTGTAGAGAAACAAATTCACACATGTTCTTTTACA 1560  
DB 188 GTAGCAAAAGCATGGCCCATGATTTTCAGAGATGGAACCTACACGATTTCTTGAGA 247  
QY 1561 CCAGAAATTCCTGGACTCCAAAGCCTGTGTATTTGTATTTATTTAGATGATGG 1620  
DB 248 AAGAGAGCGTTTGGAGTAAGATATCTGTACAGACATGATTTTGAATGCTTGG 307  
QY 1621 GAAATTTAGTGAATGTTCTTAAATTCCTGTTACGTTGTTTAAATTAAGTAAAG 1680  
DB 308 AAAAAAATGATGATTCCTTAAAGCGCTTGTATTATGCTTCCATGCTTGAACAATT 367  
QY 1681 CTATTTGGAGTAAAGTAAAGCTG 1705  
DB 368 CAAATGCTGACAAATCTCAGCG 392

## RESULT 15

US-09-071-035-81  
; Sequence 81, Application US/09071035  
; Patent No. 6448043  
; GENERAL INFORMATION:  
; APPLICANT: Gil H. Choi  
; TITLE OF INVENTION: Enterococcus faecalis Polynucleotides and Polypeptides  
; NUMBER OF SEQUENCES: 496  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Human Genome Sciences, Inc.  
; STREET: 9410 Key West Avenue  
; CITY: Rockville  
; STATE: Maryland  
; COUNTRY: USA  
; ZIP: 20850  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage  
; COMPUTER: HP Vectra 486/33  
; OPERATING SYSTEM: MSDOS version 6.2  
; SOFTWARE: ASCII Text  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/071,035  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: A. Anders Brookes  
; REGISTRATION NUMBER: 36,373  
; REFERENCE/DOCKET NUMBER: PB369P2  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (301) 309-8504  
; TELEFAX: (301) 309-8512  
; INFORMATION FOR SEQ ID NO: 81:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1680 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: linear  
US-09-071-035-81

Query Match 0.9%; Score 38.6; DB 4; Length 1680;  
Best Local Similarity 49.3%; Pred. No. 1.1;  
Matches 101; Conservative 0; Mismatches 104; Indels 0; Gaps 0;

QY 1501 GTATCCAGGAGACAGTGGTGTAGAGAAACAAATTCACACATGTTCTTTTACA 1560  
DB 286 GTAGCAAAAGCATGGCCCATGATTTTCAGAGATGGAACCTACACGATTTCTTGAGA 345  
QY 1561 CCAGAAATTCCTGGACTCCAAAGCCTGTGTATTTGTATTTATTTAGATGATGG 1620

DB 346 AAGAGAGCGTTTGGAGTAAAGATGATCCGTACAGACATGATTTGAATATGCTTGG 405  
QY 1621 GAAATTTAGTGAATGTTCTTAAATTCCTGTTACGCTGTTTAAATTAAGATAAG 1680  
DB 406 AAAAAAATGATGATTCCTTAAAGCGCTTGTATTAGCTTCTCATGCTTGAACAATT 465  
QY 1681 CTATTTGGAGTAAAGTAAAGCTG 1705  
DB 466 CAAATGCTGACAAATCTCAGCG 490

Search completed: January 16, 2004, 21:01:59  
Job time : 274 secs

Result No.	Score	Query Match	Length	DB	ID	Description
1	2584.6	60.3	3033	10	US-09-833-381-1810	Sequence 1810, App
2	1969.4	45.9	2273	12	US-10-108-2604-953	Sequence 953, App
3	730.8	17.0	1300	13	US-10-133-013-323	Sequence 223, App
4	541.6	12.6	875	13	US-10-160-162-48	Sequence 48, App
5	541.6	12.6	875	13	US-09-820-6649-48	Sequence 48, App
6	272	6.3	354	10	US-09-960-352-12867	Sequence 12867, App
7	194.4	4.5	669	13	US-10-027-632-204326	Sequence 204326, App
8	194.4	4.5	669	14	US-10-027-632-204326	Sequence 204326, App
9	130	3.0	4671	13	US-10-316-253-668	Sequence 266, App
10	130	3.0	4671	13	US-10-316-253-668	Sequence 268, App
11	126.8	3.0	4615	10	US-09-880-107-3727	Sequence 3727, App
12	126.8	3.0	4615	13	US-10-086-285-905	Sequence 405, App
13	125.2	2.9	4488	13	US-09-094-1514-9	Sequence 9, App
14	119.8	2.8	4501	12	US-10-094-686-125	Sequence 125, App
15	119.4	2.8	1140	9	US-09-764-853-362	Sequence 362, App

16	119.4	2.8	1140	11	US-09-764-876-17	Sequence 17, Appl
17	118.6	2.8	2092	11	US-09-764-876-12	Sequence 12, Appl
18	117.8	2.7	4595	10	US-09-917-8004-1531	Sequence 1531, App
19	117.4	2.7	2092	9	US-09-764-853-163	Sequence 163, App
20	116.4	2.7	2252	12	US-10-108-2608-1292	Sequence 1292, App
21	114.2	2.7	4527	13	US-09-756-247-3	Sequence 3, Appl
22	114.2	2.7	4527	13	US-09-756-247-5	Sequence 5, Appl
23	110	2.6	2041	13	US-10-292-0814-4	Sequence 4, Appl
24	110	2.6	2041	13	US-10-087-188-1	Sequence 1, Appl
25	110	2.6	4021	9	US-09-925-301-552	Sequence 552, App
26	110	2.6	4422	9	US-09-873-403-4	Sequence 4, Appl
27	110	2.6	4530	13	US-10-292-0814-2	Sequence 2, Appl
28	110	2.6	4576	13	US-10-292-0814-7	Sequence 7, Appl
29	110	2.6	4577	9	US-09-873-403-3	Sequence 3, Appl
30	110	2.6	4577	10	US-09-680-107-2236	Sequence 2236, App
31	110	2.6	4577	12	US-10-331-4964-19	Sequence 19, Appl
32	110	2.6	4577	13	US-09-873-319-408	Sequence 408, App
33	110	2.6	4577	13	US-09-960-706-654	Sequence 654, App
34	110	2.6	4577	13	US-10-292-0814-3	Sequence 3, Appl
35	110	2.6	4577	13	US-10-292-0814-5	Sequence 5, Appl
36	110	2.6	4577	14	US-10-052-817-1	Sequence 1, Appl
37	110	2.6	4577	15	US-10-076-816-56	Sequence 56, Appl
38	110	2.6	4809	13	US-10-240-965-178	Sequence 178, App
39	110	2.6	4823	13	US-09-571-392-172	Sequence 172, Appl
40	110	2.6	5092	13	US-10-006-285-404	Sequence 404, App
41	110	2.6	5092	13	US-09-571-4289-2	Sequence 2, Appl
42	106.4	2.5	4771	13	US-09-756-247-27	Sequence 27, Appl
43	99	2.3	4426	12	US-10-094-986-123	Sequence 123, App
44	93.4	2.2	4451	11	US-09-918-995-12615	Sequence 12615, App
45	85.6	2.0	5191	13	US-10-006-285-206	Sequence 206, App

## ALIGNMENTS

```

RESULT 1
US-09-833-381-1810
; Sequence 1810, Application US/09833381
; Patent No. US20020132090A1
; GENERAL INFORMATION:
; APPLICANT: Robison, Keith E.
; TITLE OF INVENTION: No. US20020132090A1 Nucleic Acid and Protein Homologs
; FILE REFERENCE: 5800-1119
; CURRENT APPLICATION NUMBER: US/09/833,381
; CURRENT FILING DATE: 2001-04-11
; PRIOR APPLICATION NUMBER: 09/516,448
; PRIOR FILING DATE: 2000-02-29
; NUMBER OF SEQ ID NOS: 2050
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1810
; LENGTH: 3033
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-833-381-1810

```

Query Match	60.3%	Score	2584.6	DB	10	Length	3033
Best Local Similarity	98.0%	Pred.	No. 0				
Matches	2648	Conservative	0	Mismatches	4	Indels	51
				Gaps			1
Qy	1636	GTCTTAAATAATCTCTGTTCAGCTTGTTTTAAATAAGATAAAGCTATATTGGAGCTAA	16995				
Db	6	GTTCGAAAAATCTCTGTTCAGCTGTTTAAAAATAAGATAAAGCTATATTGGAGTAA	65				
Qy	1696	GTGAAGCTGAACCATCTGAGAAAGTCTCTTTAGATCTCTGTACACAGCTGACTCC	1755				
Db	66	GTGAAGGTGAACCATCTGAGAAAGTCTCTTTAGATCTCTGTACACAGCTGACTCC	125				
Qy	1756	ATTATTGGAGTTAGTGTCTTGAACAAAGCTGAATCTGATGAATTCCTTAAGATTT	1815				
Db	126	ATTAGTGGATTGTAGCTGTGGAACAAAGTGTGAATCTGATGAATTCCTTAAGATTT	185				
Qy	1816	ACAATGGAATAATGTGTCATGAGTGGAACTTTATTAACAAGATATTATTAGCATG	1875				

Dh 186 ACAATGAAAATGTCGTCATGATGGAAGTGAACCTTATAACAAGATATTATTAGGCATG 245  
Qy 1876 TTCAATGAAATCTTTTGCAGCTCTTTCAGGAATGTGACCTGCGATATTGACAGATCAAAAC 1935  
Db 246 TTCAATGAAATCTTTTGCAGCTCTTTCAGGAATGTGACCTGCGATATTGACAGATCAAAAC 305  
Qy 1936 CTCAGGAAGATATTATGATGTGTATTATGACAAATGCAAAATATGCTGAGAGGTTTATG 1995  
Db 306 CTCAGGAAGATATTATGATGTGTATTATGACAAATGCAAAATATGCTGAGAGGTTTATG 365  
Qy 1996 GAGGAATATGAAAGCATATTGTATGATATATGATGCTTTCTTTGGGTAGAGTCCATAT 2055  
Db 366 GAGGAATATGAAAGCATATTGTATGATATATGATGCTTTCTTTGGGTAGAGTCCATAT 425  
Qy 2056 GTCCGAAGACATTTTCCAGAGCTTGATTTGGCTAGACACCAATGCGTTTACAGAT 2115  
Db 426 GTCCGAAGACATTTTCCAGAGCTTGATTTGGCTAGACACCAATGCGTTTACAGAT 485  
Qy 2116 TACCAAGATTTGAAGTATCTGTATCTGATCTGATCTTCTTTGGGTGCTACTGTGTTT 2175  
Db 486 TACCAAGATTTGAAGTATCTGTATCTGATCTTCTTTGGGTGCTACTGTGTTT 545  
Qy 2176 GTGATCTGAGGACCTGGGTCTTGGACTAACAATACTCCAGTGGAGCTCCAGGCTTC 2235  
Db 546 GTGATCTGAGGACCTGGGTCTTGGACTAACAATACTCCAGTGGAGCTCCAGGCTTC 605  
Qy 2236 CAACCATTTTTCATTTTGTGATCTTCCCTACTCTGTTATCAGAGGTAGAAATTTGCT 2295  
Db 606 CAACCATTTTTCATTTTGTGATCTTCCCTACTCTGTTATCAGAGGTAGAAATTTGCT 665  
Qy 2296 TTGGAATATCTATATTCATATTATTGAAAGATGCGACTGAGGTTAAGTATCATTTAG 2355  
Db 666 TTGGAATATCTATATTCATATTATTGAAAGATGCGACTGAGGTTAAGTATCATTTAG 725  
Qy 2356 AAAAGTGAATTTGATATTCTATATGATCTTCAATGAAATTAATGCGACGCGACACG 2415  
Db 726 AAAAGTGAATTTGATATTCTATATGATCTTCAATGAAATTAATGCGACGCGACACG 785  
Qy 2416 CAGACCTCTGCTGTCCTCCAGTGAAGATGCGGCACTGTTCTTTTCCATCAGGCGCA 2475  
Db 786 CAGACCTCTGCTGTCCTCCAGTGAAGATGCGGCACTGTTCTTTTCCATCAGGCGCA 845  
Qy 2476 CATCTGGGAAATTTCTATCAAGTCAAGCTCTTTTCAACCCTGCTTGTGATGCTGC 2535  
Db 846 CATCTGGGAAATTTCTATCAAGTCAAGCTCTTTTCAACCCTGCTTGTGATGCTGC 905  
Qy 2536 ACCGAGATTTTATAGTAAAGCTGGAAGAAATGAAATATCATATTCATCATCTTA 2595  
Db 906 ACCGAGATTTTATAGTAAAGCTGGAAGAAATGAAATATCATATTCATCATCTTA 965  
Qy 2596 TTAGACTGACCTGACATAGGCTACAGAGTACCTGAAACCTTTGAGTTTCTCATTTTCT 2655  
Db 966 TTAGACTGACCTGACATAGGCTACAGAGTACCTGAAACCTTTGAGTTTCTCATTTTCT 1025  
Qy 2656 CCTATATACAGTACGTCGAGTGAAGAGTTCAGATCTGCAATTTGAGATGTTCTTGCT 2715  
Db 1026 CCTATATACAGTACGTCGAGTGAAGAGTTCAGATCTGCAATTTGAGATGTTCTTGCT 1085  
Qy 2716 CCTTCATCAATGCTTACCTCATGATTTGAGATGCTTATGCTGTGTGGAACAGAAC 2775  
Db 1086 CCTTCATCAATGCTTACCTCATGATTTGAGATGCTTATGCTGTGTGGAACAGAAC 1145  
Qy 2776 ATGATTAATTTTGTCCAAATATTACATTTTGTGATTTCTGATTAATAAAGAAACAATG 2835  
Db 1146 ATGATTAATTTTGTCCAAATATTACATTTTGTGATTTCTGATTAATAAAGAAACAATG 1205  
Qy 2836 ACAGATATTGAAAGAAAGAGCTTTTCATTTATGAGGCAAGGTTACAGAGAACTT 2895  
Db 1206 ACAGATATTGAAAGAAAGAGCTTTTCATTTATGAGGCAAGGTTACAGAGAACTT 1265  
Qy 2896 CTCTATCAGAGGAAAGATGCTTTTCAGTCTTTTGGGAAATATGACCTTCTGGAGC 2955  
Db 1266 CTCTATCAGAGGAAAGATGCTTTTCAGTCTTTTGGGAAATATGACCTTCTGGAGC 1325

Qy 2956 ACTTGTTGTACGCTTTTGTTTAAGATGTTTCTTGAAAGCCGATCCTTATCATATAT 3015  
Db 1326 ACTTGTTGTACGCTTTTGTTTAAGATGTTTCTTGAAAGCCGATCCTTATCATATAT 1385  
Qy 3016 GATCAGAAATGTGTTACACAGAAATACACTTGGCTTAAAGAAATCAGAAATCCAGCT 3075  
Db 1386 GATCAGAAATGTGTTACACAGAAATACACTTGGCTTAAAGAAATCAGAAATCCAGCT 1445  
Qy 3076 GAATTTGGATCCAGGAAGATGATTCATATGAGCTTCAAGTGGCAATTAAGTCCA 3135  
Db 1446 GAATTTGGATCCAGGAAGATGATTCATATGAGCTTCAAGTGGCAATTAAGTCCA 1505  
Qy 3136 GTAAACCTTACAGCTTATTTGATCTTCTCTGCGATATTAAGAAATATCAGCTTAC 3195  
Db 1506 GTAAACCTTACAGCTTATTTGATCTTCTCTGCGATATTAAGAAATATCAGCTTAC 1565  
Qy 3196 ATGATGTGCAAGATCTATCCATTTTGGAGTCTGAATTCAGTATAGGAATTTCAAG 3255  
Db 1566 ATGATGTGCAAGATCTATCCATTTTGGAGTCTGAATTCAGTATAGGAATTTCAAG 1625  
Qy 3256 AATTATCTAGGCTTATATCTTATATGATGATGATCAGTGGGAGTCTTAAAGCGAG 3315  
Db 1626 AATTATCTAGGCTTATATCTTATATGATGATGATCAGTGGGAGTCTTAAAGCGAG 1685  
Qy 3316 GAAGCTTTGAATATGCTGACTTGGAGAGCAAGAAAGATGCGATGCAATTTCTGGGT 3375  
Db 1686 GAAGCTTTGAATATGCTGACTTGGAGAGCAAGAAAGATGCGATGCAATTTCTGGGT 1745  
Qy 3376 TCATCAGATCCAACTTTTGAATCTCTGCGAGCAGCTTCTCGATATTTGAAGTTGCA 3435  
Db 1746 TCATCAGATCCAACTTTTGAATCTCTGCGAGCAGCTTCTCGATATTTGAAGTTGCA 1805  
Qy 3436 GCGTATGCACTGCTCAGCTCTTACATTTCAATTTCAAGCTTCAAGGAAATCCCAATTATG 3495  
Db 1806 GCGTATGCACTGCTCAGCTCTTACATTTCAATTTCAAGCTTCAAGGAAATCCCAATTATG 1865  
Qy 3496 AGGTGCTAAGCAGGCAAGAAATAGCTTGGGTGTTTTCATCTACAGATACCACT 3555  
Db 1866 AGGTGCTAAGCAGGCAAGAAATAGCTTGGGTGTTTTCATCTACAGATACCACT 1925  
Qy 3556 GTGCTTTAAAGCTCTGTGTAATTTGACAGCTTATGAAATCAGAAAGACAAATATC 3615  
Db 1926 GTGCTTTAAAGCTCTGTGTAATTTGACAGCTTATGAAATCAGAAAGACAAATATC 1985  
Qy 3616 CAATGACGTTGAGGAGGCTTACCTCAAGTCTT----- 3651  
Db 1986 CAATGACGTTGAGGAGGCTTACCTCAAGTCTT----- 2045  
Qy 3652 -----CTTGTGTGTACAGCAATGCGATTAATAT 3684  
Db 2046 AACCGCTTACTCTTACAGACAGAGCTTGTGTGTACAGCAATGCGATTAATAT 2105  
Qy 3685 TCCGCAATGTTTGGATTTGCTATTTGTACCTCAATGTTGATTAATGGAAGCT 3744  
Db 2106 TCCGCAATGTTTGGATTTGCTATTTGTACCTCAATGTTGATTAATGGAAGCT 2165  
Qy 3745 TCTGGTCTTCTGAAGACGAAGATCTATGCAAAATCAAGAAACCTTTGATTAAGT 3804  
Db 2166 TCTGGTCTTCTGAAGACGAAGATCTATGCAAAATCAAGAAACCTTTGATTAAGT 2225  
Qy 3805 GCTGTAAAGAAATTAAGATGATCTCAATCATGTGATTTGAATGTGTACAGCTT 3864  
Db 2226 GCTGTAAAGAAATTAAGATGATCTCAATCATGTGATTTGAATGTGTGTACAGCTT 2285  
Qy 3865 TCGGGCCCGGATGAGGTGATGCTTATGAAAGTTAACTATTAAGTGGCTTTATG 3924  
Db 2286 TCGGGCCCGGATGAGGTGATGCTTATGAAAGTTAACTATTAAGTGGCTTTATG 2345  
Qy 3925 GTGCTTTCAAGACATTTCTGTAGAGGAGACGTGAAGAAATGGAATATGATGCA 3984  
Db 2346 GTGCTTTCAAGACATTTCTGTAGAGGAGACGTGAAGAAATGGAATATGATGCA 2405



QY 3985 AAACCTCACTCTATTAGATTCTGTAATGAACCAGTTTGTGTAAATATCTCTGCT 4044  
 DB 2406 AAACCTCACTCTATTAGATTCTGTAATGAACCAGTTTGTGTAAATATCTCTGCT 2465  
 QY 4045 GTGAGAACTTTAAGTTTCAATACCCAGATGCTTCACTGTCATAGTGAATTAATAT 4104  
 DB 2466 GTGAGAACTTTAAGTTTCAATACCCAGATGCTTCACTGTCATAGTGAATTAATAT 2525  
 QY 4105 GAGCCAGAGAGACAGCGGTGAGAGATTACAACTCTGAAGTGAAGCTTCTCTGTGAC 4164  
 DB 2526 GAGCCAGAGAGAGACAGCGGTGAGAGATTACAACTCTGAAGTGAAGCTTCTCTGTGAC 2585  
 QY 4165 CTTTGAGTGAATGTCAGAGGCTGCGCTCTTGTGAGATGAGCTTCAAGCTCCATCAT 4224  
 DB 2586 CTTTGAGTGAATGTCAGAGGCTGCGCTCTTGTGAGATGAGCTTCAAGCTCCATCAT 2645  
 QY 4225 CACTCTTCACTGATTTTATTTTCTGTTTCAAGCTTCTGACTTATATGAATCTTGGCTG 4284  
 DB 2646 CACTCTTCACTGATTTTATTTTCTGTTTCAAGCTTCTGACTTATATGAATCTTGGCTG 2705  
 QY 4285 TGA 4287  
 DB 2706 TGA 2708

RESULT 2  
 US-10-108-260A-953  
 ; Sequence 953, Application US/10108260A  
 ; Publication No. US20040005560A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: HELIX RESEARCH INSTITUTE  
 ; TITLE OF INVENTION: No. US20040005560A1e1 full length cDNA  
 ; FILE REFERENCE: H1-A0106  
 ; CURRENT APPLICATION NUMBER: US/10/108, 260A  
 ; CURRENT FILING DATE: 2002-03-27  
 ; NUMBER OF SEQ ID NOS: 5458  
 ; SOFTWARE: PatentIn Ver. 2.1  
 ; SEQ ID NO 953  
 ; LENGTH: 2273  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 US-10-108-260A-953

Query Match 45.9%; Score 1969.4; DB 12; Length 2273;  
 Best Local Similarity 99.9%; Pred. No. 0;  
 Matches 1970; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGAGAGGCGCCACCGCTCTGACCGCGCCACCTCTCTGCTGTGACCGCGCGCTG 60  
 DB .66 ATGAGAGGCGCCACCGCTCTGACCGCGCCACCTCTCTGCTGTGACCGCGCGCTG 125  
 QY 61 GCCGTGGCTCCCGGGCTCGGTTCCTGTGACAGCCCGAGGATCATAGGCCCGAGGA 120  
 DB 126 GCCGTGGCTCCCGGGCTCGGTTCCTGTGACAGCCCGAGGATCATAGGCCCGAGGA 185  
 QY 121 AATGTACTATTGGGGTGAAGCTTCTGAAACATGCGCTTCAAGATGACTGGAAGGGG 180  
 DB 186 AATGTACTATTGGGGTGAAGCTTCTGAAACATGCGCTTCAAGATGACTGGAAGGGG 245  
 QY 181 GAGCTGCTCAAGACAGCATCAACCTCACTGCTCTGTCTGTGAGACAGAGAGAGCTTT 240  
 DB 246 GAGCTGCTCAAGACAGCATCAACCTCACTGCTCTGTCTGTGAGACAGAGAGAGCTTT 305  
 QY 241 GAAAAAGGCTTTTAAAGACACTTACTTCCATCATCTGTAAGAGTGAAGTGAAG 300  
 DB 306 GAAAAAGGCTTTTAAAGACACTTACTTCCATCATCTGTAAGAGTGAAGTGAAG 365  
 QY 301 ATTATGAGCTAGCTGTAACCGGACGTACCCAGAGTGAATTTATTTCTATATAGTACC 360  
 DB 366 ATTATGAGCTAGCTGTAACCGGACGTACCCAGAGTGAATTTATTTCTATATAGTACC 425  
 QY 361 CGCTTATCAATTGAGACCAAGAAATATCTGTCTTCAATCAACAGACAGGCTTATAC 420

DB 426 CGCTTATCAATTGAGACCAAGAAATATCTGTCTTCAATCAACAGACAGGCTTATAC 485  
 QY 421 AAGCCAAAGCAGAGAGTGAAGTTCGATGTTTACACTCTTCCAGATTTTAAGCTTAC 480  
 DB 486 AAGCCAAAGCAGAGAGTGAAGTTCGATGTTTACACTCTTCCAGATTTTAAGCTTAC 545  
 QY 481 AAAACCTCTTAAACATCTCAATTAAGACCCCAATCAAAATTTGATCCAGAGTGTG 540  
 DB 546 AAAACCTCTTAAACATCTCAATTAAGACCCCAATCAAAATTTGATCCAGAGTGTG 605  
 QY 541 TCACAAACAAAGTATCTTGAGTCAATTCGAAACCTTTGACGTATCTTCCATCCATA 600  
 DB 606 TCACAAACAAAGTATCTTGAGTCAATTCGAAACCTTTGACGTATCTTCCATCCATA 665  
 QY 601 CTTGATGACTGCTCAATCAAGTTCAGTGAATGACGACATTAATCAATCATTTGAC 660  
 DB 666 CTTGATGACTGCTCAATCAAGTTCAGTGAATGACGACATTAATCAATCATTTGAC 725  
 QY 661 GTTTCAGAAATATGATTAACCAAAATTTGAAGTGAAGTTCGACAGACCATTAATGTTCT 720  
 DB 726 GTTTCAGAAATATGATTAACCAAAATTTGAAGTGAAGTTCGACAGACCATTAATGTTCT 785  
 QY 721 ATGAATTCAGATTTAAATGTAACATCAAGCAAGATATCATATGGAAGCCAGTG 780  
 DB 786 ATGAATTCAGATTTAAATGTAACATCAAGCAAGATATCATATGGAAGCCAGTG 845  
 QY 781 AAAGAGAGCTAGAGCTTACATTTTACCTTATCTTGTGGGAAAGAAATATAT 840  
 DB 846 AAAGAGAGCTAGAGCTTACATTTTACCTTATCTTGTGGGAAAGAAATATAT 905  
 QY 841 ACAAACAACTTTAAGATTAATGATCTGCAAACTTCTCTTTAATGATGAAGATGA 900  
 DB 906 ACAAACAACTTTAAGATTAATGATCTGCAAACTTCTCTTTAATGATGAAGATGA 965  
 QY 901 AATGTATGATGATCTTCAATGAGACTTCTGTAATCTGTGATCTATCTTCCCTGACCA 960  
 DB 966 AATGTATGATGATCTTCAATGAGACTTCTGTAATCTGTGATCTATCTTCCCTGACCA 1025  
 QY 961 GTAGAAATTTTAAACAGAGTGAAGATTAAGATTAATGATTAATGATTAAGCACT 1020  
 DB 1026 GTAGAAATTTTAAACAGAGTGAAGATTAAGATTAATGATTAATGATTAAGCACT 1085  
 QY 1021 AATGTCTCTTCAACCAATGATTAATGATTAATGATTAATGATTAATGATTAATGAT 1080  
 DB 1086 AATGTCTCTTCAACCAATGATTAATGATTAATGATTAATGATTAATGATTAATGAT 1145  
 QY 1081 AAGCCATCTCTCAACTTCAAGCCCACTGTGAAGTGAATCTGTGTGATGGAACCACTG 1140  
 DB 1146 AAGCCATCTCTCAACTTCAAGCCCACTGTGAAGTGAATCTGTGTGATGGAACCACTG 1205  
 QY 1141 ACTCTTGAAGAAAGAAATATATGATCTATTAAGTGAACAGAGAACTATATCTGAG 1200  
 DB 1206 ACTCTTGAAGAAAGAAATATATGATCTATTAAGTGAACAGAGAACTATATCTGAG 1265  
 QY 1201 TACTGAGAGGATCTAAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 1260  
 DB 1266 TACTGAGAGGATCTAAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 1325  
 QY 1261 ACTGTCCCCCAAGTGAACCTTTTAAAGTGAATTCCTCAATCTGTGAGAGATTCAGTGA 1320  
 DB 1326 ACTGTCCCCCAAGTGAACCTTTTAAAGTGAATTCCTCAATCTGTGAGAGATTCAGTGA 1385  
 QY 1321 CTACAGTTGAAGGCTATTTCTGTGTGATTAAGATTAAGATTAAGATTAAGATTAAG 1380  
 DB 1386 CTACAGTTGAAGGCTATTTCTGTGTGATTAAGATTAAGATTAAGATTAAGATTAAG 1445  
 QY 1381 AAGTCTCTAGTAAGACATCAATCAATCAATCAATCAATCAATCAATCAATCAATCA 1440  
 DB 1446 AAGTCTCTAGTAAGACATCAATCAATCAATCAATCAATCAATCAATCAATCAATCA 1505  
 QY 1441 TCGCTTTTGAAGT 1500  
 DB 1506 TCGCTTTTGAAGT 1565

QY 1501 GTATCCAGGGGACAGTTGGTGGCTGTAGAGAAACAAATTCACATGTTCTCTTTACA 1560  
DB 1566 GTATCCAGGGGACAGTTGGTGGCTGTAGAGAAACAAATTCACATGTTCTCTTTACA 1625  
QY 1561 CCAGAAATTTCTGGACCTCCAAAAGCCCTGTATTTGCTATTTATTTGAAGATGATGGG 1620  
DB 1626 CCAGAAATTTCTGGACCTCCAAAAGCCCTGTATTTGCTATTTATTTGAAGATGATGGG 1685  
QY 1621 GAAATTTATAGTGTCTCTAAAAATTCCTGTTCAGCTGTGTTTTTAAAAATTAAGTAAAG 1680  
DB 1686 GAAATTTATAGTGTCTCTAAAAATTCCTGTTCAGCTGTGTTTTTAAAAATTAAGTAAAG 1745  
QY 1681 CTATTTAGAGTAAAGTAAAGCTGAAACATCTGAGAAAGTCTCTTACAGATCTCTGTG 1740  
DB 1746 CTATTTAGAGTAAAGTAAAGCTGAAACATCTGAGAAAGTCTCTTACAGATCTCTGTG 1805  
QY 1741 ACACAGCTGACTCATATGTTGGGATTTAGCTGTGCAAAAAGTGTGATCTGATGAAT 1800  
DB 1806 ACACAGCTGACTCATATGTTGGGATTTAGCTGTGCAAAAAGTGTGATCTGATGAAT 1865  
QY 1801 GCCTCTATATGATTTATACAAATGGAATGTCATGAGTTGGAATTTATTAACACAGGA 1860  
DB 1866 GCCTCTATATGATTTATACAAATGGAATGTCATGAGTTGGAATTTATTAACACAGGA 1925  
QY 1861 TATTTATAGGCAATGTTGATGAATTTCTTTGCACTTTTCAAGATGTGACTGAGGTA 1920  
DB 1926 TATTTATAGGCAATGTTGATGAATTTCTTTGCACTTTTCAAGATGTGACTGAGGTA 1985  
QY 1921 TTGACAGATGCAAACTTCACAGAGATTAATTTAGTGTGTTATGACAAAT 1971  
DB 1986 TTGACAGATGCAAACTTCACAGAGATTAATTTAGTGTGTTATGACAAAT 2036

RESULT 3  
US-10-133-013-223  
; Sequence 223, Application US/10133013  
; Publication No. US20030166903A1  
; GENERAL INFORMATION:  
; APPLICANT: Astromoff, Anna  
; APPLICANT: Bandman, Olga  
; APPLICANT: Cocks, Benjamin G.  
; TITLE OF INVENTION: GENES ASSOCIATED WITH VASCULAR DISEASE  
; FILE REFERENCE: PA-0049 US  
; CURRENT APPLICATION NUMBER: US/10/133,013  
; PRIORITY FILING DATE: 2002-04-25  
; PRIORITY APPLICATION NUMBER: 60/287,067  
; PRIORITY FILING DATE: 2001-04-27  
; NUMBER OF SEQ ID NOS: 271  
; SOFTWARE: PERL Program  
; SEQ ID NO 223  
; LENGTH: 1300  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc feature  
; OTHER INFORMATION: Incyte ID No. US20030166903A1 3658034CB1  
US-10-133-013-223

Query Match 17.0%; Score 730.8; DB 13; Length 1300;  
Best Local Similarity 93.6%; Pred. No. 4.8e-179;  
Matches 805; Conservative 0; Mismatches 2; Indels 53; Gaps 2;  
QY 3479 AGGAAATCCCAATTTATGAGTGGCTTAAGCAGGCAAAAGAAATGCTTGGGTGTTTGCAT 3538  
DB 1 AGGAAATCCCAATTTATGAGTGGCTTAAGCAGGCAAAAGAAATGCTTGGGTGTTTGCAT 60  
QY 3539 CTACTCAGATACACATGCTGGCTTTAAAGGCTGTGCTGAATTTGACGCCCTATATGATA 3598  
DB 61 CTACTCAGATACACATGCTGGCTTTAAAGG--CTGTCTGAATTTGACGCCCTATATGATA 118  
QY 3599 CAGAAAGCAAAATATCCAACTGACCGTGAAGGAGCCCTAGCTCAACAAGTCT----- 3651

DB 119 CAGAAAGCAAAATATCCAACTGACCGTGAAGGAGCCCTAGCTCAACAAGTCTGTAAGT 178  
QY 3652 -----CTGCTGTGTACAGC 3667  
DB 179 TTCTGATGACACACACACACCGCTTACTCTTGACAGACAGACCTTGCTGTGTACAGC 238  
QY 3668 CAATGCAATTAATTAATTTCCGCAAAATGTTTGAATTTGCTATTTGTCACTCAATGTTG 3127  
DB 239 CAACGCAATTAATTAATTTCCGCAAAATGTTTGAATTTGCTATTTGCTCACTCAATGTTG 298  
QY 3728 TATTAATGTAAGGCTTCTGGGCTCTTCTGAAAGACGAATATCTATCCAAATCAAGAAG 3787  
DB 299 TATTAATGTAAGGCTTCTGGGCTCTTCTGAAAGACGAATATCTATCCAAATCAAGAAG 358  
QY 3788 CTTTGAATTTAGATGTTGCTGTAAAGAAATTAAGATGATCTCAATCATGTGGATTTGA 3847  
DB 359 CTTTGAATTTAGATGTTGCTGTAAAGAAATTAAGATGATCTCAATCATGTGGATTTGA 418  
QY 3848 ATGTGTACAAACCTTTTGGGCCCCGGGTAGAGTGCATGAGCTTTATGAAAGTTAAC 3907  
DB 419 ATGTGTACAAACCTTTTGGGCCCCGGGTAGAGTGCATGAGCTTTATGAAAGTTAAC 478  
QY 3908 TATTAAGTGGCTTTATGAGTGCCTTCAAGACATTTCTCTGACGAGACAGTGAAGAAAG 3967  
DB 479 TATTAAGTGGCTTTATGAGTGCCTTCAAGACATTTCTCTGACGAGACAGTGAAGAAAG 538  
QY 3968 TGGAAATGATCATGGAAGAACTCAACCTTATTTAGATTTGTAATGAAGAACAGTTT 4027  
DB 539 TGGAAATGATCATGGAAGAACTCAACCTTATTTAGATTTGTAATGAAGAACAGTTT 598  
QY 4028 GTGTAAATATTTCTGCTGTGAGAACTTTAAAGTTTCAATTAACCAAGATGTTCAAGTGT 4087  
DB 599 GTGTAAATATTTCTGCTGTGAGAACTTTAAAGTTTCAATTAACCAAGATGTTCAAGTGT 658  
QY 4088 CCATAGTGCATTTCTATAGACCAAGAGACAGGCGGTGAGAACTTCAACTCTGAAGTGA 4147  
DB 659 CCATAGTGCATTTCTATAGACCAAGAGACAGGCGGTGAGAACTTCAACTCTGAAGTGA 718  
QY 4148 AGCTGTCTCTGCTGAGACCTTTGACAGTATGTCAGAGGCTGCGCTCTTGTGAGAGTGA 4207  
DB 719 AGCTGTCTCTGCTGAGACCTTTGACAGTATGTCAGAGGCTGCGCTCTTGTGAGAGTGA 778  
QY 4208 CTTGAGGCTCCATCATCATCTTCACTCTTCACTCACTTTTATTTTCTGTTCAAGCTTCTGACT 4267  
DB 779 CTTGAGGCTCCATCATCATCTTCACTCTTCACTCTTATTTTCTGTTCAAGCTTCTGACT 838  
QY 4268 TTATGAACTTTGGCTGTGA 4287  
DB 839 TTATGAACTTTGGCTGTGA 858

RESULT 4  
US-10-160-162-48  
; Sequence 48, Application US/10160162  
; Publication No. US20030166541A1  
; GENERAL INFORMATION:  
; APPLICANT: Ruben et al.  
; TITLE OF INVENTION: 83 Human Secreted Proteins  
; FILE REFERENCE: P2012P2  
; CURRENT APPLICATION NUMBER: US/10/160,162  
; PRIORITY FILING DATE: 2002-06-04  
; PRIORITY APPLICATION NUMBER: 60/295,558  
; PRIORITY FILING DATE: 2001-06-05  
; PRIORITY APPLICATION NUMBER: 09/236,557  
; PRIORITY FILING DATE: 1999-01-26  
; PRIORITY APPLICATION NUMBER: PCT/US98/15949  
; PRIORITY FILING DATE: 1998-07-29  
; PRIORITY APPLICATION NUMBER: 60/054,212  
; PRIORITY FILING DATE: 1997-07-30  
; PRIORITY APPLICATION NUMBER: 60/054,209  
; PRIORITY FILING DATE: 1997-07-30  
; PRIORITY APPLICATION NUMBER: 60/054,234  
; PRIORITY FILING DATE: 1997-07-30

```

/ PRIOR APPLICATION NUMBER: 60/054,218
/ PRIOR FILING DATE: 1997-07-30
/ PRIOR APPLICATION NUMBER: 60/054,214
/ PRIOR FILING DATE: 1997-07-30
/ PRIOR APPLICATION NUMBER: 60/054,236
/ PRIOR FILING DATE: 1997-07-30
/ PRIOR APPLICATION NUMBER: 60/054,215
/ PRIOR FILING DATE: 1997-07-30
/ PRIOR APPLICATION NUMBER: 60/054,211
/ PRIOR FILING DATE: 1997-07-30
/ PRIOR APPLICATION NUMBER: 60/054,217
/ PRIOR FILING DATE: 1997-07-30
/ PRIOR APPLICATION NUMBER: 60/054,213
/ PRIOR FILING DATE: 1997-07-30
/ PRIOR APPLICATION NUMBER: 60/055,968
/ PRIOR FILING DATE: 1997-08-18
/ PRIOR APPLICATION NUMBER: 60/055,969
/ PRIOR FILING DATE: 1997-08-18
/ PRIOR APPLICATION NUMBER: 60/055,972
/ PRIOR FILING DATE: 1997-08-18
/ PRIOR APPLICATION NUMBER: 60/056,561
/ PRIOR FILING DATE: 1997-08-19
/ PRIOR APPLICATION NUMBER: 60/056,534
/ PRIOR FILING DATE: 1997-08-19
/ PRIOR APPLICATION NUMBER: 60/056,729
/ PRIOR FILING DATE: 1997-08-19
/ PRIOR APPLICATION NUMBER: 60/056,543
/ PRIOR FILING DATE: 1997-08-19
/ PRIOR APPLICATION NUMBER: 60/056,727
/ PRIOR FILING DATE: 1997-08-19
/ PRIOR APPLICATION NUMBER: 60/056,554
/ PRIOR FILING DATE: 1997-08-19
/ PRIOR APPLICATION NUMBER: 60/056,730
/ PRIOR FILING DATE: 1997-08-19
/ NUMBER OF SEQ ID NOS: 353
/ SOFTWARE: Patentin Ver. 2.0
/ SEQ ID NO 48
/ LENGTH: 875
/ TYPE: DNA
/ ORGANISM: Homo sapiens
US-10-160-162-48

Query Match      12.6%; Score 541.6; DB 13; Length 875;
Best Local Similarity 99.8%; Pred. No. 6.4e-130;
Matches 541; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 3746 CTGGCTCTTCTGAGAGCAAGATCTATCCAAATCAAGAGCTTTGATTTAGATGTTG 3805
DB 15 CTGGCTCTTCTGAGAGCAAGATCTATCCAAATCAAGAGCTTTGATTTAGATGTTG 74
QY 3806 CTGTAAGAAATAAAGATGATCTCAATCATGTGATTTGAATGCTCAAGCTTTT 3865
DB 75 CTGTAAGAAATAAAGATGATCTCAATCATGTGATTTGAATGCTCAAGCTTTT 134
QY 3866 CGGGCCGGGTAGAGTGCATGCTCTTATGGAAGTTAACTTAACTGCTTTATG 3925
DB 135 CGGGCCGGGTAGAGTGCATGCTCTTATGGAAGTTAACTTAACTGCTTTATG 194
QY 3926 TGCCTTCAGAGCAATTTCTCTGAGCGAGACAGTGAAGAAAGTGAATATGATCGAA 3985
DB 195 TGCCTTCAGAGCAATTTCTCTGAGCGAGACAGTGAAGAAAGTGAATATGATCGAA 254
QY 3986 AACTCAACTCTATTAGATTCTGTAATGAAGAACCAAGTTTGTGTAATATTCCTGCG 4045
DB 255 AACTCAACTCTATTAGATTCTGTAATGAAGAACCAAGTTTGTGTAATATTCCTGCG 314
QY 4046 TGAGAACTTTAAAGTTTCAAAATACCAAGATGCTCAGTGCATAGTGAATTAAGT 4105
DB 315 TGAGAACTTTAAAGTTTCAAAATACCAAGATGCTCAGTGCATAGTGAATTAAGT 374
QY 4106 AGCCAAAGAGACAGGGGTGAGAGTTAACTTGAAGTGAAGCTGTCTCTGTGACC 4165
DB 375 AGCCAAAGAGACAGGGGTGAGAGTTAACTTGAAGTGAAGCTGTCTCTGTGACC 434
```

```

QY 4166 TTGCGATGATGTCAGGGCTGCCCTCTTGAGAGATGAGCTTCAGGCTCCATCATC 4225
DB 435 TTGCGATGATGTCAGGGCTGCCCTCTTGAGAGATGAGCTTCAGGCTCCATCATC 494
QY 4226 ACTCTTCAGTACTTTTATTTCTGTTTCAAGCTTCTGACTTTATGGAACCTTGCGCT 4285
DB 495 AMTCTTCAGTACTTTTATTTCTGTTTCAAGCTTCTGACTTTATGGAACCTTGCGCT 554
QY 4286 GA 4287
DB 555 GA 556

RESULT 5
US-09-820-649-48
/ Sequence 48, Application US/09820649
/ Publication No. US20030199683A1
/ GENERAL INFORMATION:
/ APPLICANT: Ruben et al.
/ FILE OF INVENTION: 83 Human Secreted Proteins
/ FILE REFERENCE: P2012P1
/ CURRENT APPLICATION NUMBER: US/09/820,649
/ CURRENT FILING DATE: 2001-03-30
/ PRIOR APPLICATION NUMBER: US/09/236,557
/ PRIOR FILING DATE: 1999-01-26
/ PRIOR APPLICATION NUMBER: PCT/US98/15949
/ PRIOR FILING DATE: 1998-07-29
/ PRIOR APPLICATION NUMBER: 60/054,212
/ PRIOR FILING DATE: 1997-07-30
/ PRIOR APPLICATION NUMBER: 60/054,209
/ PRIOR FILING DATE: 1997-07-30
/ PRIOR APPLICATION NUMBER: 60/054,234
/ PRIOR FILING DATE: 1997-07-30
/ PRIOR APPLICATION NUMBER: 60/054,218
/ PRIOR FILING DATE: 1997-07-30
/ PRIOR APPLICATION NUMBER: 60/054,214
/ PRIOR FILING DATE: 1997-07-30
/ PRIOR APPLICATION NUMBER: 60/054,236
/ PRIOR FILING DATE: 1997-07-30
/ PRIOR APPLICATION NUMBER: 60/054,215
/ PRIOR FILING DATE: 1997-07-30
/ PRIOR APPLICATION NUMBER: 60/054,211
/ PRIOR FILING DATE: 1997-07-30
/ Remaining Prior Application data removed - See File Wrapper or PALM.
/ NUMBER OF SEQ ID NOS: 353
/ SOFTWARE: Patentin Ver. 2.0
/ SEQ ID NO 48
/ LENGTH: 875
/ TYPE: DNA
/ ORGANISM: Homo sapiens
US-09-820-649-48

Query Match      12.6%; Score 541.6; DB 13; Length 875;
Best Local Similarity 99.8%; Pred. No. 6.4e-130;
Matches 541; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 3746 CTGGCTCTTCTGAGAGCAAGATCTATCCAAATCAAGAGCTTTGATTTAGATGTTG 3805
DB 15 CTGGCTCTTCTGAGAGCAAGATCTATCCAAATCAAGAGCTTTGATTTAGATGTTG 74
QY 3806 CTGTAAGAAATAAAGATGATCTCAATCATGTGATTTGAATGCTCAAGCTTTT 3865
DB 75 CTGTAAGAAATAAAGATGATCTCAATCATGTGATTTGAATGCTCAAGCTTTT 134
QY 3866 CGGGCCGGGTAGAGTGCATGCTCTTATGGAAGTTAACTTAACTGCTTTATG 3925
DB 135 CGGGCCGGGTAGAGTGCATGCTCTTATGGAAGTTAACTTAACTGCTTTATG 194
QY 3926 TGCCTTCAGAGCAATTTCTCTGAGCGAGACAGTGAAGAAAGTGAATATGATCGAA 3985
DB 195 TGCCTTCAGAGCAATTTCTCTGAGCGAGACAGTGAAGAAAGTGAATATGATCGAA 254
```

QY 3986 AACTCAACCTCTATTAGATTCTGTAAATGAACCCAGTTTGTGTAAATATTCCTGCTG 4045  
DB 255 AACTCAACCTCTATTAGATTCTGTAAATGAACCCAGTTTGTGTAAATATTCCTGCTG 314  
QY 4046 TGAGAACTTTAAAGTTTCAAAATGCCAAGATGCTTCAGTGCATATGTGATTAATCATG 4105  
DB 315 TGAGAACTTTAAAGTTTCAAAATGCCAAGATGCTTCAGTGCATATGTGATTAATCATG 374  
QY 4106 AGCCAAAGAGACAGCGCTGAGAAAGTTACAACTGAGAGAGCTGCTCTGTGACC 4165  
DB 375 AGCCAAAGAGACAGCGCTGAGAAAGTTACAACTGAGAGAGCTGCTCTGTGACC 434  
QY 4166 TTTGAGATGATGCCAGGCTGCTGCTGCTGAGATGAGAGCTTCAGGCTCCATCATC 4225  
DB 435 TTTGAGATGATGCCAGGCTGCTGCTGCTGAGATGAGAGCTTCAGGCTCCATCATC 494  
QY 4226 ACTCTGAGATTTTATTTTCTGTTCAAGCTTCCTGATGATGAACTTTGGGCTGT 4285  
DB 495 AMCTTCAGTCACTTTTATTTTCTGTTCAAGCTTCCTGATGATGAACTTTGGGCTGT 554  
QY 4286 GA 4287  
DB 555 GA 556

## RESULT 6

US-09-960-352-12867  
; Sequence 12867, Application US/09960352  
; Patent No. US20020137139A1  
; GENERAL INFORMATION:  
; APPLICANT: Warren, Wesley C.  
; APPLICANT: Tao, Nengping  
; APPLICANT: Byatt, John C.  
; APPLICANT: Mathialagan, Nagappan  
; TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND  
; FILE REFERENCE: 16511.006/37-21(10298)C  
; CURRENT APPLICATION NUMBER: US/09/960.352  
; CURRENT FILING DATE: 2001-09-24  
; NUMBER OF SEQ ID NOS: 15112  
; SEQ ID NO 12867  
; LENGTH: 354  
; TYPE: DNA  
; ORGANISM: Bos taurus  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: (285)  
; OTHER INFORMATION: unsure at all n locations  
; OTHER INFORMATION: Clone ID: 55-LIH3057-014-01-K1-F12  
US-09-960-352-12867

Query Match 6.3%; Score 272; DB 10; Length 354;

Best Local Similarity 89.1%; Pred. No. 5, 1e-60; Mismatches 36; Indels 0; Gaps 0;

QY 2084 TTTGCTGAGACCAACATGGGTTTACAGATTTCACAAAGATTGGAAGTAATCTACTG 2143  
DB 26 TTGGCTTGGCCCCCACTGGGTTTCCAGATTTCACAAAGATTGGAAGTAATCTACTG 85  
QY 2144 ATTCTATCATCTTTGGGCTGCTACTGTTTGTGATCTGAGAGACTGGGCTTGGAC 2203  
DB 86 ATTCTATCATCTTTGGGCTGCTACTGTTTGTGATCTGAGAGATTAGGATCTTGGAC 145  
QY 2204 TAAACAACATCCGAGTGGAGTCCCAAGCTTCCAAACATTTTTCATTTTTCATCTC 2263  
DB 146 TAAACAACATCCGAGTGGAGTCCCAAGCTTCCAAACATTTTTCATTTTTCATCTC 205  
QY 2264 CCTACTCTGTATGAGAGGTGAAGATTGGCTTGGAAATAGTATATTCATATTTGA 2323  
DB 206 CATACTCTGTATGAGAGGTGAAGATTGGCTTGGAAATAGTATATTCATATTTGA 265  
QY 2324 AAGATGCCACTGAGGTTAAGTAAATCATTTGAGAAAAGTGACAAATTTGATTTGAATGA 2383  
DB 2383 AAGATGCCACTGAGGTTAAGTAAATCATTTGAGAAAAGTGACAAATTTGATTTGAATGA 2383

DB 266 AAGATGCCACTGAGGTTAAGTAAATCATTTGAGAAAAGTGATGCTTTGATATTTCAATG 325  
QY 2384 CTTCAGTGAATTAATGACAGGAC 2412  
DB 326 CTTCAGTGAATTAATGACAGGAC 354

## RESULT 7

US-10-027-632-204326/c  
; Sequence 204326, Application US/10027632  
; Publication No. US20030204075A9  
; GENERAL INFORMATION:  
; APPLICANT: Wang, David G.  
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide  
; FILE REFERENCE: 108827.129  
; CURRENT APPLICATION NUMBER: US/10/027.632  
; CURRENT FILING DATE: 2002-04-30  
; PRIOR APPLICATION NUMBER: US 60/218,006  
; PRIOR FILING DATE: 2000-07-12  
; PRIOR APPLICATION NUMBER: US 60/198,676  
; PRIOR FILING DATE: 2000-04-20  
; PRIOR APPLICATION NUMBER: US 60/193,483  
; PRIOR FILING DATE: 2000-03-29  
; PRIOR APPLICATION NUMBER: US 60/185,218  
; PRIOR FILING DATE: 2000-02-24  
; PRIOR APPLICATION NUMBER: US 60/167,363  
; PRIOR FILING DATE: 1999-11-23  
; PRIOR APPLICATION NUMBER: US 60/156,358  
; PRIOR FILING DATE: 1999-09-28  
; PRIOR APPLICATION NUMBER: US 60/146,002  
; PRIOR FILING DATE: 1999-08-09  
; NUMBER OF SEQ ID NOS: 325720  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 204326  
; LENGTH: 689  
; TYPE: DNA  
; ORGANISM: Human  
US-10-027-632-204326

Query Match 4.5%; Score 194.4; DB 13; Length 689;

Best Local Similarity 99.5%; Pred. No. 1, 3e-39; Mismatches 1; Indels 0; Gaps 0;

QY 3352 GAAGTGGCATGGAATTTGGGTTGTCATCAGAGTCCAACTTTCTGACTCTGGACCCA 3411  
DB 485 GAAGTGGCATGGAATTTGGGTTGTCATCAGAGTCCAACTTTCTGACTCTGGACCCA 426  
QY 3412 CGCTCCCTGGATATGGAAGTTGAGGCTATGACCTGCTCTACACTTTTCAATTTGAG 3471  
DB 425 CGCTCCCTGGATATGGAAGTTGAGGCTATGACCTGCTCTACACTTTTCAATTTGAG 366  
QY 3472 ACTTTCGAGGATTTCCAAATTATGAGTGGCTTAAGCAGGAAGAAATGCTGGGTGT 3531  
DB 365 ACTTTCGAGGATTTCCAAATTATGAGTGGCTTAAGCAGGAAGAAATGCTGGGTGT 306  
QY 3532 TTTGATCTACTCAGG 3547  
DB 305 TTTGATCTACTCAGG 290

## RESULT 8

US-10-027-632-204326/c  
; Sequence 204326, Application US/10027632  
; GENERAL INFORMATION:  
; APPLICANT: Wang, David G.  
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide  
; FILE REFERENCE: 108827.129  
; CURRENT APPLICATION NUMBER: US/10/027.632  
; CURRENT FILING DATE: 2002-04-30  
; PRIOR APPLICATION NUMBER: US 60/218,006  
; PRIOR FILING DATE: 2000-07-12

PRIOR APPLICATION NUMBER: US 60/198,676  
 PRIOR FILING DATE: 2000-04-20  
 PRIOR APPLICATION NUMBER: US 60/193,483  
 PRIOR FILING DATE: 2000-03-29  
 PRIOR APPLICATION NUMBER: US 60/185,218  
 PRIOR FILING DATE: 2000-02-24  
 PRIOR APPLICATION NUMBER: US 60/167,363  
 PRIOR FILING DATE: 1999-11-23  
 PRIOR APPLICATION NUMBER: US 60/156,358  
 PRIOR FILING DATE: 1999-09-28  
 PRIOR APPLICATION NUMBER: US 60/146,002  
 PRIOR FILING DATE: 1999-08-09  
 NUMBER OF SEQ ID NOS: 325720  
 SOFTWARE: FastSeq for Windows Version 4.0  
 SEQ ID NO 204326  
 LENGTH: 689  
 TYPE: DNA  
 ORGANISM: Human  
 US-10-027-632-204326

Query Match 4.5%; Score 194.4; DB 14; Length 689;  
 Best Local Similarity 99.5%; Pred. No. 1.3e-39;  
 Matches 195; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3352 GAAGGTGCATGCAATTCGGGTGTCATCAGAGTCCAACTTTCTGACTCTGGAGCCA 3411  
 DB 485 GCAAGGTGCATGCAATTCGGGTGTCATCAGAGTCCAACTTTCTGACTCTGGAGCCA 426  
 QY 3412 GCGTCCCTGGATATGAAAGTTGACGCTTATGACCTGCTTCACTTTTACAAATTTTCG 3471  
 DB 425 GCGTCCCTGGATATGAAAGTTGACGCTTATGACCTGCTTCACTTTTACAAATTTTCG 366  
 QY 3472 ACTTCGAGGGAATCCCAATTAATGAGGTGCTAAGAGGCAAGAAATAGTTGGGTGCT 3531  
 DB 365 ACTTCGAGGGAATCCCAATTAATGAGGTGCTAAGAGGCAAGAAATAGTTGGGTGCT 306  
 QY 3532 TTTCGATCTACTACG 3547  
 DB 305 TTTCGATCTACTACG 290

RESULT 9  
 US-10-316-253-266  
 Sequence 266, Application US/10316253  
 Publication No. US20030162706A1  
 GENERAL INFORMATION:  
 APPLICANT: The Procter & Gamble Company  
 APPLICANT: Peters, Kevin  
 APPLICANT: Thompson, Larry  
 APPLICANT: Wang, Feng  
 APPLICANT: Greis, Kenneth  
 TITLE OF INVENTION: Angiogenesis Modulating Proteins  
 FILE REFERENCE: 8865M  
 CURRENT APPLICATION NUMBER: US/10/316,253  
 CURRENT FILING DATE: 2002-12-10  
 PRIOR APPLICATION NUMBER: US 60/355,295  
 PRIOR FILING DATE: 2002-02-08  
 NUMBER OF SEQ ID NOS: 308  
 SOFTWARE: PatentIn version 3.1  
 SEQ ID NO 266  
 LENGTH: 4677  
 TYPE: DNA  
 ORGANISM: Rattus norvegicus  
 FEATURE:  
 NAME/KEY: CDS  
 LOCATION: (52)..(4554)  
 OTHER INFORMATION:  
 US-10-316-253-266

Query Match 3.0%; Score 130; DB 13; Length 4677;  
 Best Local Similarity 47.3%; Pred. No. 3e-22;  
 Matches 575; Conservative 0; Mismatches 590; Indels 51; Gaps 4;

QY 2056 GTCCGAAGACATTTTCCAGAGACTTGGATTTGGCTAGACACCAACATGGGTTTACAGATT 2115  
 DB 2284 GTCCGAAGACATTTTCCAGAGACTTGGATTTGGCTAGACACCAACATGGGTTTACAGATT 2343  
 QY 2116 TACCAAGAAATTTGAAGTAACTGTAACCTGATTTCTATCACTTCTGTGGTGGCTACTGGTTT 2175  
 DB 2344 GACGGTGAAGTCCCGTAAGAGTCCCTACACCAATCACTGAGTGAAGGCCAGTCAATTC 2403  
 QY 2176 GTGATCTGAGAGACCTGGGCTTTGAGCTAACTCACTCCAGTGGAGCTCCAGACCTTC 2235  
 DB 2404 TGCTGTCTGAGACACTGCGGCTTGTCTCTCTCAACA---TCTCCATTAATCTTTC 2460  
 QY 2236 CAACCAATTTTCAATTTTGAATCTTCTCACTCTGTTATCAGAGGTGAAGAAATTTGCT 2295  
 DB 2461 CAGCCCTTCTCTGAGACCTCACTCTCCCTACTGTGTGTTGAGGGAAGCAATTTATTC 2520  
 QY 2236 TTGGAATTAATCTATTTTCAATTTTGAAGATGCACTGAGGTTAAGTAATCTTTAG 2355  
 DB 2521 CTCAAGCTACCGTACCTCACTCACTGCTCACTGATTCGATTCGATTCGATTCGATTCG 2580  
 QY 2356 AAA-----AGTCAAAATTTGATTTCTAATGACT 2385  
 DB 2581 ATGTCTCTGATTTCTCTGAGAGTCCCTGTGGAGGCCATGAAGATTTCTAATGATCTGT 2640  
 QY 2386 TCAAGTGAATTAATGATGCAAGGCCAGGCAAGCCCTTCTGTGTTCCAGTGAAGATGGG 2445  
 DB 2641 GGAATTAAGAAAGAAACCGTGTCTGAGGCTGTGACCCCAAGTGTGGGGAGGTGAAC 2700  
 QY 2446 GCAATGTTCTTTTCCATCAGGCCAACAATCTGGAGAAATTTCTATCACTGACACA 2505  
 DB 2701 TTCAAGCTACTGCAAGAGCTTTGCACTCTCAGAACTGTGTGCAAAATGAGTGCAGAA 2760  
 QY 2506 GCTCTTCAACCACTGCTTCTGATGCTGTCACCCAGATGATTTTGAAGGCTGAAGA 2565  
 DB 2761 GTGCGACCTTTGTAAGAGGACACTGTAGTGAAGCCCGTATATGAGCTTAAGGA 2820  
 QY 2566 ATGAAATTAATTAATTTCAATTCATCTTATTAAGTGAAGTGAAGTGAAGT 2625  
 DB 2821 ATGAAATTAATTAATTTCAATTCATCTTATTAAGTGAAGTGAAGTGAAGT 2876  
 QY 2626 ACCCTGAATCTTGAAGTTCTCATTTCTCTTAATACAGTGAAGTGAAGTGAAGT 2685  
 DB 2877 -----GGAATTAATTAATTTCAATTCATCTTATTAAGTGAAGTGAAGTGAAGT 2931  
 QY 2686 CAGATCACTGCAATTTGAGATGTTTGTGTCTTCTCAATCAATGCTTGAAGCTCAATTTGAT 2745  
 DB 2932 ACACAGTCCGTTTGGGGAATTAATTAATTTGAGTGAAGTGAAGTGAAGTGAAGT 2991  
 QY 2746 CGATGCTTATGCTGTGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 2805  
 DB 2992 CAGATGCTTATGCTGTGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 3051  
 QY 2806 TTGATTAATCTGACTTAAAGAAAGAAACCACTGACAGTAAATTTGAAGAAAGCTCTTCA 2865  
 DB 3052 CTGAGATTAATCAATGAAGACACAGAGTGAAGAGGAGTCAAGTCAAGGCAATTTGAG 3111  
 QY 2866 TTATGAGGCAAGTACCAGAGAGATTTCTTATCAGAGGAAAGTGTCTTTTCACT 2925  
 DB 3112 TACCTCATCAGCGGTGACAGAGGAGCTGAATCAATTCAGACAGTGAAGTGTCTATCAGC 3171  
 QY 2926 GCTTTGGGAATTAATG-----ACCTTCTGAGGACACTGTGTCAGCTTTGTT 2976  
 DB 3172 ACATTCGGGACCGGATGATGAGGACAGTCAAGGAAACATTTGCTCACTGCAATTTGTG 3231  
 QY 2977 TTAAGATGTTTCTTGAAGCCGATCTTACATGATTAATTTGAATGATGATTTACACAGA 3036  
 DB 3232 CTCAAGGCTTCTGCTCAAGCTCAAGTCAATTAATTAATTAATTAATTAATTAATTAAT 3291  
 QY 3037 ACATTAATCTGCTTAAAGGACATCAAGAAATCAAGGATGATTTTGGAGTCCAGAGAA 3096  
 DB 3292 GCTTCAATTTGCTCTGATGAAGAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3351  
 QY 3097 GTGATTAATGATGAGCTTCAAGGTGCAATTAAGTCAAGTCAAGTCAAGTCAAGTCAAGT 3156

Accession	Sequence	Position
Db	3352 CTGCTCAACAAATGCGATAGAGGGTGTGTGATGATGAAGGACACTTCTTGCTTAATTC	3411
Qy	3157 GTAACCTTCTCTCGTGGATATAGAAAGTACAGCCTAAcATTGATGTGCAAGCTATAC	3216
Db	3412 ACCATTGCTCTGTGCTGAGATGCGCCCTGCTGTCACTACAGTGTGTGTAAGCTCTA	3471
Qy	3217 CATTTTGGAGTCTG 3232	
Db	3472 TTCTGCTGTGAACGG 3487	

```

RESULT 10
US-10-316-253-268
Sequence 268, Application US/10316253
Publication No. US20030162706A1
GENERAL INFORMATION:
APPLICANT: The Procter & Gamble Company
APPLICANT: Peters, Kevin
APPLICANT: Thompson, Larry
APPLICANT: Wang, Feng
APPLICANT: Greis, Kenneth
TITLE OF INVENTION: Angiogenesis Modulating Proteins
FILE REFERENCE: 8865M
CURRENT APPLICATION NUMBER: US/10/316,253
CURRENT FILING DATE: 2002-12-10
PRIOR APPLICATION NUMBER: US 60/355,295
PRIOR FILING DATE: 2002-02-08
NUMBER OF SEQ ID NOS: 108
SOFTWARE: PatentIn version 3.1
SEQ ID NO 268
LENGTH: 4701
TYPE: DNA
ORGANISM: Rattus norvegicus
FEATURE:
NAME/KEY: CDS
LOCATION: (95)..(4597)
OTHER INFORMATION:
US-10-316-253-268

```

Query Match	3.0%;	Score 130;	DB 13;	Length 4701;
Best Local Similarity	47.3%;	Pred. No. 3e-22;		
Matches 575;	Conservative	0;	Mismatches 590;	Indels 51;
				Gaps 4;

OY	2056	TTCCGAAGCAATTTTCCAGAGCTTGATTTGGCTGACACCAACATGGTTTACAGATT	2115
Db	2327	GTGCGGAAGTACTTCCCGAAAACCTGGATCTGGGACATGTGTCCGCTGACCTTATCCGGT	2386
OY	2116	TACCAAGAATTGAGTAACGTACGTATCTATCACTCTTGGGTGGCTACTGCTTTT	2175
Db	2387	GACGTGATGTGCCGTTAAAGTCTCCCTGACACATCACTGATGTGAAGCCAGTCATTC	2446
OY	2176	GTGATCTCTGAGAGCCTGGGTCTTGGACTAACAACTACTCCAGTGGAGCTCCAAAGCTTC	2235
Db	2447	TGCTTGTCTGGAACACTACTGCGCTTGTGTCTCTCCACACA--TCTCCCATTAAGTCTTC	2503
OY	2236	CAACCAATTTTCATTTTCTTTTGGATCTTCCCTACTCTGTATTCAGAGGTGAAGAAATTTGCT	2295
Db	2504	CAGCCCTCTTCTCTGAGCTCACTTCCCTTACTGTGTGTTCGAGCGGAAGCAATTTATC	2563
OY	2296	TTGGAATTAATCATATATTCATTTTGGAAAGATGCACTGAGGTTAAGTATCATTTAG	2355
Db	2564	CTCAAAAGTACCGTACTCACTACTACTGCTCACTGCAATTCGATGCAATGTAGCCTAAG	2623
OY	2356	AAA-----AGTGAACAATTTGATATTTCTAATGACT	2385
Db	2624	ATGTCTCTGATTTCTCGCAGATCCCTGTGGGAGACCAATGAAGATTTCAATGCAATCGT	2683
OY	2386	TCAAGTGAATTAATGATCCACAGGCCACAGCAGACCTTTCTGTGTTCCAGTGAAGATGGG	2445
Db	2684	GGAATATGAAGAAAACCGTGTCTCTGGGCTGTGACCCCAAGATGCTGGGGGAGGTGAAC	2743
OY	2446	GCAACTGTCTTTTCCATCAGGCCAACACATCTGGGAGAAATTTCTATCACTACACTCA	2505

Db	2744	TTCAAGCTCTACTGCAGAAAGCTTTTGCAGTCTTCACAGAACTGTGTGCCAATTAAAGTGGACGAA	2803
OY	2506	GCTCTTTTCACCCACTGTGCTTGTGATGCTGTCAACCAGATGATTTTAAAGGCTGAAGA	2565
Db	2804	GTGCACGCCCTTGTACAGAAAGGACACTGTATGTGAAGCCCGTAATATGTGAGCCTGAAGGA	2863
OY	2566	ATGAAAAATCATATTACAACTCCATCTTTATTAGACTTTGACTGACAAATAGGCTTACAGT	2625
Db	2864	ATTGAGAAGAGCAAAAGTACAAACACTGTATTGCCCCAAGATGTGTGATTACA----	2919
OY	2626	ACCTGAAACCTTTGAGTTTCTCATTTCCCTCTTAATACAGTACTGSCAGTGAAGAAGTT	2685
Db	2920	-----GGAAAATTGGAACACTGATCTTCCAGCCCATGTGTTGAAGATCTGCCAGGGCT	2974
OY	2686	CAGATCACTGCAATTGAGATGTTCCTTGCTCTTCCATCAATGGCTTAGCCTCAATTGATT	2745
Db	2975	ACACAGTCCGTTTGGGCGATATATCTAGGCTCTGCGATGCAAAACCTCCAGAAATCTTCTC	3034
OY	2746	CGGATGCCCTTATGGCTGTGTGAAACAGAACATGATTAATTTTGCTCCAAATATTACATT	2805
Db	3035	CAGATGCCCTTATGGCTGTGGGGAAACAAACATGTCTCTTCGCTCCCTTAACATTCAGTT	3094
OY	2806	TTGATATTATCTGATTAAGAAACAACTGACAGATTAATTGAAAGAAAGACCTCTTCA	2865
Db	3095	CTGAGATATCTCAATGAGACACAGCAGCTGACAGGCGCATCAAGTCCAAAGCCATTAGC	3154
OY	2866	TTTATGAGGCAAGGTTACCAAGAGAACTTCTATTCAGAGGGAGAGATGGCTCTTTTCAGT	2925
Db	3155	TACCTCATCAGCGGGGTACAGAGGCACTGAACATATAGCACAGTGAAGGTTACTAATACGC	3214
OY	2926	GCTTTGGGAATTATG-----ACCTTCTGGGAGCACTTGTTGTACAGCTTTGTT	2976
Db	3215	ACATTCCGGGGACCGCGGTATAGGCAAGTCAAGGGAACCTTGGCTCACCTGCACTTTGTG	3274
OY	2977	TTAAGATGTTTCCCTTGAAGCCGATCCCTTACATAGATTTGATCGAATGTGTTACACAGA	3036
Db	3275	CTCAAGGCTTTGCTCAAGCTCACTCAATCTATATAGAAAAAGACACATACAAAT	3334
OY	3037	ACATPACACTTGGCTTAAAGACATCAGAAATCCCAACGCGTAATTTTGGATCCACGAGA	3096
Db	3335	GCTTTCATTTGGCTCTCTGATGAAACAAAGGAGAACGTTGTTTCCAAAGCTCGAATCC	3394
OY	3097	GTCATCATAGTAGCTTCAAAGTGCGCAATPAAAGTCCAGTACACTTAACGCTTATTT	3156
Db	3395	CTGCTCAACATGAGTAGAAGGGGGTGTGTGATGATGAAGTACACTCTGCGCTTATATC	3454
OY	3157	GTAACCTCTCTCCCTGGGATATAGAAATATACGCTTAAACATTAATGTGCAAGACTTATC	3216
Db	3455	ACCAATGCTCTGCTGGAGATGCCCTTCCTGTCACTACAGTGTGTTGTTAATGCTCTA	3514
OY	3217	CATTTTGGAGTCTG 3232	
Db	3515	TTCTGCTGTGAAACGG 3530	

US-09-880-107-3727  
 : Sequence 3727, Application US/09880107  
 : Patent No. US20020142981A1  
 : GENERAL INFORMATION:  
 : APPLICANT: Horne, Darci T.  
 : APPLICANT: Vockley, Joseph G.  
 : APPLICANT: Schert, Uwe  
 : APPLICANT: Gene Logic, Inc..  
 : TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer  
 : FILE REFERENCE: 44921-5028-WO  
 : CURRENT APPLICATION NUMBER: US/09/880,107  
 : CURRENT FILING DATE: 2001-06-14  
 : PRIOR APPLICATION NUMBER: US 60/211,379  
 : PRIOR FILING DATE: 2000-06-14  
 : PRIOR APPLICATION NUMBER: US 60/237,054  
 : PRIOR FILING DATE: 2000-10-02



```

RESULT 12
US-10-006-265-405
; Sequence 405, Application US/10006285
; Publication NO. US2003016585A1
; GENERAL INFORMATION:
; APPLICANT: Mary Jane Cunningham
; APPLICANT: Matthew R. Kaser
; TITLE OF INVENTION: MARKER GENES RESPONDING TO TREATMENT WITH TOXINS
; FILE REFERENCE: PA-0039 US
; CURRENT APPLICATION NUMBER: US/10/006,285
; CURRENT FILING DATE: 2001-12-05
; NUMBER OF SEQ ID NOS: 314
; SOFTWARE: PERL Program
; SEQ ID NO 405

```

```

RESULT 13
US-09-981-151A-9
; Sequence 9, Application US/09981151A
; Publication No. US20030212256A1
; GENERAL INFORMATION:
; APPLICANT: Edinger, Salomte R
; APPLICANT: Gerlach, Valerie
; APPLICANT: MacDougall, John R
; APPLICANT: Malvanchar, Muriel M
; APPLICANT: Smithson, Glenda
; APPLICANT: Millet, Isabelle
; APPLICANT: Peyman, John A
; APPLICANT: Stone, David J
; APPLICANT: Gunther, Erik
; APPLICANT: Ellerman, Karen
; APPLICANT: Shinkets, Richard A

```



APPLICANT: Gorman, Linda  
APPLICANT: Malyankar, Uriel M.  
APPLICANT: Boldog, Ferenc  
APPLICANT: Guo, Xiaojia  
APPLICANT: Shenoy, Suresh  
APPLICANT: Padigaru, Muralidhara  
APPLICANT: Taupier, Raymond J., Jr.  
APPLICANT: Miller, Charles  
APPLICANT: Casman, Stacie  
APPLICANT: Pena, Carol  
APPLICANT: Gangolli, Esma  
APPLICANT: Gusev, Vladimir  
APPLICANT: Smithson, Glennda  
APPLICANT: zeinusen, Bryan  
APPLICANT: Gerlach, Valerie  
APPLICANT: Pochart, Pascal  
APPLICANT: Fernandes, Elma  
APPLICANT: Shinkets, Richard  
APPLICANT: Rastelli, Luca  
APPLICANT: Spaderma, Steven  
APPLICANT: Larocheille, William  
APPLICANT: Zhong, Mei  
TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME, AND METHOD  
FILE REFERENCE: 21402-290 B  
CURRENT APPLICATION NUMBER: US/10/094,886  
CURRENT FILING DATE: 2002-03-07  
PRIOR APPLICATION NUMBER: 60/274,322  
PRIOR FILING DATE: 2001-03-08  
PRIOR APPLICATION NUMBER: 60/313,182  
PRIOR FILING DATE: 2001-08-17  
PRIOR APPLICATION NUMBER: 60/288,052  
PRIOR FILING DATE: 2001-05-02  
PRIOR APPLICATION NUMBER: 60/318,510  
PRIOR FILING DATE: 2001-09-10  
PRIOR APPLICATION NUMBER: 60/274,281  
PRIOR FILING DATE: 2001-03-08  
PRIOR APPLICATION NUMBER: 60/314,018  
PRIOR FILING DATE: 2001-08-21  
PRIOR APPLICATION NUMBER: 60/274,194  
PRIOR FILING DATE: 2001-03-08  
PRIOR APPLICATION NUMBER: 60/274,849  
PRIOR FILING DATE: 2001-03-09  
PRIOR APPLICATION NUMBER: 60/296,693  
PRIOR FILING DATE: 2001-06-07  
PRIOR APPLICATION NUMBER: 60/313,626  
PRIOR FILING DATE: 2001-08-21  
Remaining Prior Application data removed - See File Wrapper or PALM.  
NUMBER OF SEQ ID NOS: 298  
SOFTWARE: PatentIn 2.1  
SEQ ID NO 125  
LENGTH: 4501  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (1)..(4309)  
US-10-094-886-125  
Query Match 2.8%; Score 119.8; DB 12; Length 4501;  
Best Local Similarity 45.7%; Pred. No. 1.3e-19;  
Matches 596; Conservative 0; Mismatches 682; Indels 27; Gaps 4;  
QY 2040 GGATGACATCCACATGTCGGAAGCATTTCCAGAGCATTTGGCTGAGACGAA 2099  
DB 2088 GGCTATATTCACAAAGTAAGAACAACTTCCAGAGCATGATGTGGACCTGCTGAG 2147  
QY 2100 CATGGGTTACAGGATTTACCAAGATTTGAAGTAAGTAAGTCTGATCTATGACTTTG 2159  
DB 2148 TGTGATTCCTCAGGCTCGCCAACTTTCCTTCCTCCTGATGATGATGATGATGATG 2207  
QY 2160 GGTGCTACTGTTTGTGATCTCTGAGAGCTGGGCTTTGATCTAACAAGTCTGACT 2219  
DB 2208 GGAGGCAAGTGGCTTTGTGTGATGTGATGACGTTGATTTGGCATTTCTCTAACAAC 2267

QY 2220 GGAGCTCCAGGCTTCCAAACATTTTTCATTTTGTGATCTTCCCTACTCTGTATGAG 2279  
DB 2268 T---CTAGAGCTCCCAACCTTCTTATGATGTGCTACACCTTTTGGTGTGCA 2324  
QY 2280 AGGTGAAGATTTGCTTTGGAATTAATGATTAATTAATTTGGAAGATGCACTGAGGT 2339  
DB 2325 AAATGACAAATTTGATTTGATTTGATGATGATGATGATGATGATGATGATGATGAT 2384  
QY 2340 TAAAGTATCATGTGGAAGAAAGTGAATTTGATTTCTAATGATCTTAAGTGAATTA 2399  
DB 2385 TTCTGTCAAGTGAAGAGTCTCAGAAATTTGAAGCAAAATTTCAATCTTGAATTA 2444  
QY 2400 TCCAGAGGCCACCGAGACCTCTGCTTCCAGTGAAGATGGGGCACTGTTCTTT 2459  
DB 2445 TGGCAGTGAAGTATTTCAAGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2504  
QY 2460 TCCCATCAGGCCAACACATCTGGGAGAAATTCCT-ATCAAGTCAAGCTCTTTGACCCA 2518  
DB 2505 GAAATTTGGGTAAAGATATATCACTGATGTTGCTGATGCTCAAAACAGAGTGTGCC 2564  
QY 2519 CTGCTTCAAGTCTGTACCCAGATGATTTTATGTAAGGCTGAAGGATTAAGAAATCAT 2578  
DB 2565 AAATGAAGAAATGAGCAGCAAAAGCTAAACTGAAAGACATGCTGTCCAAAGCTTCTT 2624  
QY 2579 ATTCAATCCATCTTATTAGACTTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 2638  
DB 2625 AGTAAAGCTGAAGATTTGAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2684  
QY 2639 TG-----AGTTCTCATTTCTCTATATGACGAGTGAAGTGAAGTGAAGTGAAGTGAAG 2678  
DB 2685 TCCCAAGCTTCCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2744  
QY 2679 AAGATTTGATCATCTGCAATTTGAGATGTTCTTGTCTTCTCATCAATGCTTGAAGCTC 2738  
DB 2745 CAGAGGCTTTTCACTGTGTGGGGGATTTCTAGAGATTTGCTTGCAGAAATCTG3--T 2801  
QY 2759 ATTGATGGAATGCTTATGAGTGTGAGTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2798  
DB 2802 TGTTCCTCAAAATGCTTATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2861  
QY 2799 TTACATTTTGGATTTATCTGACTAATAAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2858  
DB 2862 TTATGTTTGGAGCTATGGAATATCAAGAGGAGTGAAGAGGAGGAGGAGGAGGAGGAGG 2921  
QY 2859 TCTTTCAATTTAGAGGCAAGTTTACAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2918  
DB 2922 TTTCTTCTCTTATCTAATGTTATCAAAAGGCAATTTCTTCAAAAACCTGATGATG 2981  
QY 2919 TTTGAGGCTTTGGGAATTTATGACCTTTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2978  
DB 2982 CTATAGTGTGTTTGGAGCAGAGTCAAGAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3041  
QY 2979 AAGATGTTCTTGAAGCCGATCTTATCATGATATGATGATGATGATGATGATGATGATG 3038  
DB 3042 TAAAGCATTTGAGAGAAATGAATAATATGTTATGATGATGATGATGATGATGATGATG 3101  
QY 3039 ATCACTGGCTTAAAGGAGTCAAGAAATCCAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3098  
DB 3102 CTATATGCTGCTTCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3161  
QY 3099 GATTCATGATGAGCTTCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3158  
DB 3162 TTTCAACACGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3221  
QY 3159 AACTTCTCTCTGAGATATGAAGATATCAGCTTAACATGATGAGGAGGAGGAGGAGGAG 3218  
DB 3222 TGGGATGTTCTTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3281  
QY 3219 TTTTGTGAGTGTGAATTCAGTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3278  
DB 3282 TTGCTTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3341





PUBMED REFERENCE AUTHORS	11042159 3 Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Kono, H., Akiyama, J., Nishi, K., Katsuna, T., Tashiro, H., Itoh, M., Sun, I., Iehi, Y., Nakamura, S., Hazama, M., Nishino, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwara, S., Inoue, K., Ozawa, Y., Izawa, M., Ohara, E., Watabiki, M., Yoneda, Y., Ishikawa, T., Toga, K., Kato, K., Matsuda, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y. RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer Genome Res. 10 (11), 1757-1771 (2000)
JOURNAL MEDLINE PUBMED REFERENCE AUTHORS	20530913 11076861 4 Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y., Aizawa, T., Hara, A., Fukunishi, Y., Kono, H., Adachi, J., Fukuda, S., Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamada, I., Saito, T., Okazaki, Y., Gojobori, T., Bono, H., Kasukawa, T., Saito, R., Kadota, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T., Fleischmann, W., Gaasterland, T., Gissi, C., King, B., Kochiwa, H., Kuehl, P., Lewis, S., Matsuo, Y., Nikaido, I., Perle, G., Quekeshub, J., Schriml, L. M., Staudt, F., Suzuki, R., Tomita, M., Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H., Baldarelli, R., Barsh, G., Blake, J., Boffelli, D., Bojunga, N., Carninci, P., de Bonaldo, M. F., Brownstein, M. J., Bult, C., Fleischer, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D., Hofmann, M., Hume, D. A., Kamiya, M., Lee, N. H., Lyons, P., Marchionni, L., Mashima, J., Mazzarelli, J., Mombereis, P., Nordone, P., Rising, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H., Sato, K., Schonbach, C., Seta, T., Shibata, Y., Storch, K. F., Suzuki, H., Toyooka, K., Wang, K. H., Weitz, C., Whitaker, C., Wilming, L., Wynshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawai, H., Kohlsch, S. and Hayashizaki, Y. Functional annotation of a full-length mouse cDNA collection Nature 409 (6821), 685-690 (2001)
JOURNAL MEDLINE PUBMED REFERENCE AUTHORS	21085660 11217851 5 The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team. Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs Nature 420, 563-573 (2002)
JOURNAL REFERENCE AUTHORS	6 (bases 1 to 4485) Adachi, J., Aizawa, K., Akimura, T., Arawaka, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hirooka, T., Hirozane, T., Hori, F., Imotani, K., Iehi, Y., Itoh, M., Kagawa, I., Kasukawa, T., Kato, H., Kawai, J., Kojima, Y., Kondo, S., Kono, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N., Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shitaki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaki-Akahira, S., Takeda, Y., Tanaka, T., Tomaru, A., Toy, T., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y. Direct Submission Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gs.riken.go.jp, URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216) cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in RIKEN Genomic Sciences Center and Genome Science Laboratory in RIKEN Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. Please visit our web site for further details. URL: http://genome.gsc.riken.go.jp/ URL: http://fantom.gsc.riken.go.jp/.

FEATURES	Location/Qualifiers
source	1. 4485 /organism="Mus musculus" /mol_type="mRNA" /strain="C57BL/6J" /db_xref="FANTOM_DB:4831440K17" /db_xref="taxon:10090" /clone="4831440K17" /issue_type="head" /clone_lib="RIKEN full-length enriched mouse cDNA library" /div_stage="0 day neonate"
misc_feature	112. 4439 /note="putative weakly similar to alpha-2-macroglobulin (fragments) [Limulus polyphemus] (PIR A36260, evidence: FASTY, 51.8%ID, 76.7%length, match=168)"
BASE COUNT	1277 a 970 c 1010 g 1228 t
ORIGIN	
Query Match	63.2%; Score 2708.8; DB 11; Length 4485;
Best Local Similarity	77.9%; Pred. No. 0;
Matches 3393; Conservative	0; Mismatches 872; Indels 91; Gaps 7;
1	ATGCAAGGCGCCACCGCTCTGACCGCCGACCTCTCTGCTGTCACCGCCGCTG 60
112	ATGCGCAGTGGAGGCTCTGACCGCGCGCCACCTCTCTGCTGTCGCGGCTG 171
61	GCGGTGCTCCGCGGCTGCTTCTGCTGTCAGCGCCGAGATCATACGCCGAGCA 120
172	GCC--GCCGAGGTCCCTTCTGCTGTCAGCGCCGAGATCATACGCCGAGCA 228
121	AATGTGACTTGGGGTGAAGCTTCTGGAACACTGCTTCTGACAGGTGATGGAAGCG 180
229	AATGTGACTTGGGGTGAAGCTTCTGGAACACTGCTTCTGACAGGTGATGGAAGCG 288
181	GAGTGTCAAGACAGATCAAACTCTCTCTGCTGTCGAGAGCAAGAGTCTT 240
289	CAGGTGTTCAAGATAGCTTCAACAAATCAGATTCATCTGGAACGAGAGAGTCTT 348
241	GAAGAGGCTCTTTAAGACACTTACCTTCCATCACTACCTCTGGAACAGTGAAG 300
349	CACGAGGCAATTTCAAGACTCTGCTTCTGCGGCACTCTCTGACAGTGAAG 408
301	ATTATGAGTACGCTGTAACCGGACGATCCAGATGAGATTTATTTCTTAATGATCC 360
409	ATTATGAGTACGCTGTAACCGGACGATCCAGATGAGATTTATTTCTTAATGATCC 468
361	CGTTATCATTTGAGACCAAGAAATATCTGTTTCAATCAACAGCAAGGCTTATAC 420
469	CGTTAATCATTTGAGACCAAGAAATATCTGTTTCAATCAACAGCAAGGCTTATAC 528
421	AAGCGAAGCAAGAGGAGTTCGATGTTACACTCTTCCAGATTTAAGCCTTAC 480
529	AAGCGAAGCAAGAGGAGTTCGATGTTACACTCTTCCAGATTTAAGCCTTAT 588
481	AAAACCTCTTAAACATTTCTATTAAAGACCCCAATCAATTTGATTCGAAGCTGTTG 540
589	AGAATCATTCCTGATCATTTTCAATTAAGATCCGAATGATGATTCGAAGCTGTTG 648
541	TCAACAAAGATGATTTGAGATCATTTTCAAAATTTTCAAGTATCTTCCATCAATA 600
649	TCTCAGAAAGTATCTTGAAGTCTTCCAAATTTTCAATCTTCAATCTTCAATCAATA 708
601	CTTGATGATCTGATCTATTCAGATTTCAAGTGAAGCAAGCAATATATCAATTTTCA 660
709	TTTGATGATCTGATCTATTCAGATTTCAAGTGAAGCAAGCAATATATCAATTTTCA 768
661	GTTTCAATATGATTTATCAAAATTTGAAGTATCTTGAAGCAAGCAATATATTTGTTCT 720
769	GTTTCAATATGATTTATCAAAATTTGAAGTATCTTGAAGCAAGCAATATATTTGTTCC 828
721	ATGAAATTAAGATTTAAATGATATCAAGCAAGCAAGTATATATGGAAGCAAGT 780



Db 829 TTGAAGCTTAAACGTTAAATGGCTCCGTCATAGCAAAAGTATACATACGGGAAGCCAGTGG 888  
Qy 781 AAGAGAGAGTAAACGTTACATTTTACCTTTATCCCTTTGGGGAAGAAAGAAATATAT 840  
Db 889 AAGGGAAGTCTGTCACCTTACATTTTATCCCTTTGGGGAAGAAAGAAATATATC 948  
Qy 841 AAAAAAATTTAAGTAAATGATGTGCAAACTTCTCTTTTAAATGATGAAGATGAA 900  
Db 949 AAAAAAATTTGAGATTAATGATTTGCAAACTTCTCTTTTGAATTAATGATGAA 1008  
Qy 901 AATGTAAT-----GGAATCTTCAATGSACTTCTGTAATACCTGAT 942  
Db 1009 AAGGTAATGAATTTGAAACCCGTAACGATGTCGGAAGGAGTTATGAAATGTGGAC 1068  
Qy 943 CTATCTTCCCTGGAACCACTAGAAATTTTAAACCAAGTGAAGAAATCACTTACAGTAT 1002  
Db 1069 CGGCTCTTCTGGGCGGGGAGAGTATAGCTACGTGACGGAATCACTCACAGTATC 1128  
Qy 1003 TCAAGAAATGTAGCACTAATGTGTTCTTCAAGCAACATGATTAATCATGATGATTTT 1062  
Db 1129 TCAAGATGGCAAGCAACAAATGTGTTTCAAGCAACATGATTAATCATGATGATTTT 1188  
Qy 1063 GATTAATCACTGTCTGGAAGCATCTCAACTTCAAGCACTGTCAGTGAAGTAACTGCT 1122  
Db 1189 GATTAATGAGTGTCTGGAAGCATCTCAACTTCAAGCACTGTCAGTGAAGTAACTGCTG 1248  
Qy 1123 GCTGATGGCAACCACTGACTCTTGAAGAAAGAAATATGATGATCACTTAAACAGTACA 1182  
Db 1249 TCTGATGGCAATCACTAATCTCTGGAAGAAATGAAATGATTTGTCACAGTATGACC 1308  
Qy 1183 CAGAGAACTAATCTGAGTACGAGCGGATCTAACAGTGAATCAGAAAATGGAAGCT 1242  
Db 1309 CAGGGAATAATTAACCACTGAGAGTCAGA-----GAGATCAGAAATGATTAAT 1358  
Qy 1243 GTTCAAAATTAATTAATCTGTCCCAAGAGTGAACCTTTTAAGATTTATCCCAATC 1302  
Db 1359 ATCCAGAACCGTAATTAATTAATCTGTCCCAAGAGTGAACCTTTTAAGATTTATCCCAATC 1418  
Qy 1303 CTGAGAGATTTCCAGTGAAGCTTGAAGGCTTATTTCTGATGATTAAGTATGATG 1362  
Db 1419 ATGTCACTTTCCGGGAGCTTACAACTGAAGGCTTATTTCTGATGATTAAGTATGATG 1478  
Qy 1363 GCAGTTCATAGTCTGTTTAAGTCTCTAGTAAAGACATACATCAATCAATCAATCAAGAT 1422  
Db 1479 ACAGTCCACAGTATGTTCACTCTCTAGTAAAGATGATCACTCAAGTTAAACAGAGAT 1538  
Qy 1423 GAAATATTAAGTGGAGTCCCTTTTGAAGTGTGATTAAGTGAACAACAGATTAAG 1482  
Db 1539 GAATATTAAGTGGAGTCCCTTTTGAAGTGTGATTAAGTGAACAACAGATTAAG 1598  
Qy 1483 GATTAAGCTATAGTATGATTCAGGGGACAGTTGTGGCTGTAGGAAACAAATATCA 1542  
Db 1599 GACTTAAGCTATAGTATTAATCAAGGACAGTTGTGGCTGTAGGAAACAAATATCA 1658  
Qy 1543 ACAATGTTCTCTTAAACACAGAAATTTCTTGAAGTCAAAAGGCTGTGATTTGAT 1602  
Db 1659 ABAATCTTCTCTTAAACACAGAGGCTTCTGGGCTCCAAAGGCTGTATTAATCCATAT 1718  
Qy 1603 TATATTAAGATGATGGGAAATTAATAGTATGTTTAAATTAATCTGTTCAAGCTTGT 1662  
Db 1719 TATATTAAGATGATGGGAAATTAATAGTATGTTTAAATTAATCTGTTCAAGCTTGT 1778  
Qy 1663 TTTAAATTAAGATTAAGTATTAATGATTAAGTAAAGTGAAGCCATCTGAGAAAGTC 1722  
Db 1779 TTTGAAATTAAGATTAAGTATTAATGATTAAGTAAAGTGAAGCCATCTGATTAAGTGC 1838  
Qy 1723 TCTTATGATCTCTGTGACACAGCTGATCTCATAGTGGATGATGATGATGATGACAA 1782  
Db 1839 TCTCTCAGATCTCTGACACAGCTGATCTCTGATGGGATGATGATGATGATGACAA 1898  
Qy 1783 AGTGAATCTGATGAATGCTCTTAATGATTAATTAATGATGAAGAAATGTGCTCATGATGG 1842  
Db 1899 AGTGAATCTGATGAAGAAATCTCAACAGTATTAATTAATGATGAAGAAATGTGCTCATGATGG 1958

Qy 1843 GAACTTTAATCAACAGATTAATTAATTAAGATGATGATTAATTTTGGCAGTCTTTCAG 1902  
Db 1959 GAACTTTAATCAACAGATTAATTAATTAAGATGATGATTAATTTTGGCAGTCTTTCAG 2018  
Qy 1903 GAATGATGATCTGGGATTTGACAGATGAAACCTCAAGAAATTAATTAATGATGATG 1962  
Db 2019 GAATGATGATCTGGGATTTGACAGATGAAACCTCAAGAAATTAATTAATGATGATG 2078  
Qy 1963 TATGATTAAGATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2022  
Db 2079 TATGATTAAGATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2138  
Qy 2023 ATTCAATGATTTTCTTGGGATGACATGATCCCAAGATTTTCCAGAGATTTG 2082  
Db 2139 TTTGAAGATGATTTCTTCAAGTATGATGATGATGATGATGATGATGATGATGATG 2198  
Qy 2083 ATTGGCTAAGACACCAATGATGATGATGATGATGATGATGATGATGATGATGATG 2142  
Db 2199 ATTGGCTAAGACACCAATGATGATGATGATGATGATGATGATGATGATGATGATG 2258  
Qy 2143 GATTCATCACTTCTGGGATGATGATGATGATGATGATGATGATGATGATGATGATG 2202  
Db 2259 GATTCATCACTTCTGGGATGATGATGATGATGATGATGATGATGATGATGATGATG 2318  
Qy 2203 CTAACTAATCTCCAGTGAAGTCCCAAGCTTCCCAAGATTTTCAATTTTGAATCTT 2262  
Db 2319 TTAACAACTGTTCCAGAGAGTGAAGATGATGATGATGATGATGATGATGATGATG 2378  
Qy 2263 CCTACTGTTTATCAGAGTGAAGATTTCTTGAAGATTAATTAATTAATTAATTTG 2322  
Db 2379 CCTACTGTTTATCAGAGTGAAGATTTCTTGAAGATTAATTAATTAATTAATTTG 2438  
Qy 2323 AAAATGCACTGAGTAAAGTATCAATTAAGAAATGCAATTAATTAATTAATTTG 2382  
Db 2439 AAAATGCACTGAGTAAAGTATCAATTAAGAAATGCAATTAATTAATTAATTTG 2498  
Qy 2383 ACTTAAGTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTTG 2442  
Db 2499 ACTTAAGTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTTG 2558  
Qy 2443 GGGGCACTGTTCTTTTCCATCAGGCAACATCTGAGGAGAAATTTCTATCACTG 2502  
Db 2559 GGGGCACTGTTCTTTTCCATCAGGCAACATCTGAGGAGAAATTTCTATCACTG 2618  
Qy 2503 ACAGTCTTTTCAACCACTGCTTCTGATGCTGACCCAGATGATTTTATTAAGCTGAA 2562  
Db 2619 ACAGTCTTTTCAACCACTGCTTCTGATGCTGACCCAGATGATTTTATTAAGCTGAA 2678  
Qy 2563 GGAATGAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTTG 2622  
Db 2679 GGAATGAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTTG 2738  
Qy 2623 AGTACCTGAAATCTTGAATTTCTATTTCTCTAATACAGTGAAGTGAAGAA 2682  
Db 2739 AGTACCTGAAATCTTGAATTTCTATTTCTCTAATACAGTGAAGTGAAGAA 2798  
Qy 2683 GTTCAATCACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2742  
Db 2799 GTTCAATCACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2858  
Qy 2743 ATTGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2802  
Db 2859 ATTGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2918  
Qy 2803 ATTGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2862  
Db 2919 ATTGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2978  
Qy 2863 TCAATTAAGGCAAGGTTACCAAGAGATTTCTATCAAGAGGAAATGCTCTTTC 2922  
Db 2979 TCAATTAAGGCAAGGTTACCAAGAGATTTCTATCAAGAGGAAATGCTCTTTC 3038

Oy	2923	AGTGCTTTGGGAATTAAGACCCCTCTGGAGACATCTGGTGTGAGCTTTGGTTTAA	2982
Dp	3039	AGTGCTTTGGGGACATTGACCTCTTCTGGAGACCTTGGCTGTGACGATTTGGTTTAA	3098
Oy	2983	TGTTTCCCTGGAAGCCGATCTTACATAGATATTTGATCGAAATGTGTTTACAGACATAC	3042
Dp	3099	TGCTTTCTGGAAGCTGATTTACTATATATGATATTTGATTCAGGATGTGTTCACGAACTAT	3158
Oy	3043	ACTTGGCTTAAAGGACATCAGAAATCCAAAGGTGAATTTTGGGATCCAGGAAGTGT	3102
Dp	3159	ACTTGGCTTAAATGCACTAAGAAATTCATGTGTATTTTGGAGCCGGAAGAGTGTAT	3218
Oy	3103	CATGTGAGCTTCAAGGTGGGCAATAAAGTCAGTAACACTTACGCGCTATATGTACT	3162
Dp	3219	CACAGTAGACTTCAAGGTGGGCACCAAAGCCGAGTAACATTACCGGCTATATTTGTACT	3278
Oy	3163	TCTCTCTGGGATATAGAAAGTATCAGCCTTAACATTGATGTGTCAAGAGCTATCCATTT	3222
Dp	3279	TCTGTCTGGGATATCAAAAGATATAGGCTATATTCGATGTATCAAGACTCAATCAAGTT	3338
Oy	3223	TTGGAAGTCTGAATTCATGAGAGAAATTTCAACAATTATTACTTATGCCCTTATTAATTAT	3282
Dp	3339	TTGGAAATTTGAATTCAGCAGAGAAATTTCCGACAAATTAATCTTAAGCAATTAATCTTAT	3398
Oy	3283	GCAATTTGTCATAGTGGGGAGTCCCTAACCAGGAAGGCTTTGAATGCTGACTTGGAGA	3342
Dp	3399	GCCCTGTCCACAGTCGGGAGCCCTTAACCAAGAGGAGCTTTGGAATTTGTAATGACGGA	3458
Oy	3343	GCAGAACAGAAAGGTGGCATCAATTTGGGCTGATCAGAGTCCAACTTTCTGACTCC	3402
Dp	3459	TCAGAAAGGAAGAGACACACAGTCTGTGTGTTCATACAGGACCTGGCGCTCTGGGTCC	3518
Oy	3403	TGGCAGCCACGCTCCCTGGATATTTGAAGTTCAGGCTTATGCACTGTCTCAACTTCTTA	3462
Dp	3519	TGGCAGCGCGCCTCCGTGGACATCGAAATTCGCTGTATGCACTGTGGCCACACTGTG	3578
Oy	3463	CAATTTTGAGACTTCTGAGAGGAATCCCATATATAGAGTGGCTTAAGAGGCAAGAAATAGC	3522
Dp	3579	CA---CAATGTGTCTGAGGGAATCCCGGTATATGAGTGTGCTATCCACCAAGAAACGC	3639
Oy	3523	TTGGGTGGTTTGGCATCTTACTCAGGATACCACTGGTGGCTTTAAAGGCTCTGTCTGAATTT	3582
Dp	3636	CTGGGAAGTTTGTATTCACCTCAGATATCTGTGTGGCCCTGGAAGCATTAATCTGAATTT	3699
Oy	3583	GCAGCCCTTAATGAATACAGAAAGCAAAATATCCAAAGTACCGTGACGGGCGCTTAAGCTCA	3642
Dp	3696	TCAGCCCTTGTGCAACAGAAATAATCAGATATCCAACTGACCGTCAAGGGCCCGGCAATC	3755
Oy	3643	CCAACTCT-----	3651
Dp	3756	CCCAAGATTCATACACTTCGAAATGACCTCCAGAACTCTTTCTCTTCAACCAAGAAAG	3815
Oy	3652	CTTGTGTGTGTACAGCCAAATGAGCAAGTTAATTTTCGCAAAATGGTTTGAATTTGTAAT	3711
Dp	3816	CTTATATGGGCTAGATCCATACCGTTAATGTTTCTGGCGAATGGCTCGGGAATTTGCTAAT	3875
Oy	3712	TGTCAAGCTCAATGTGTATATATATATGTGAAGCTTCTGGGTCTTCTAGAAGACGAAGATCT	3771
Dp	3876	TGCCAGCTTAATGTGTAATAATAGTGAAGGTTCAGGTTCTTCTTAAGACGGAGATCT	3935
Oy	3772	ATCCAAATCAAGAAGCCTTTGATTAGATGTTGCTGTAAAGAAATTAAGATGATCTC	3831
Dp	3936	ACCGAAATCAAGAAGTTTGTGATTTTAAACGTCAATGTGA--ATTAAGAGAGACGACATTT	3992
Oy	3832	AATCATGTGGAATTTGAATGTGTGTATCAAGCTTTTCGGGCGCGGGTAAGAGTGGCATGGCT	3891
Dp	3993	AGTCACTGAAATCTGAATGTGTGTGCAAGTCACTTGGGTTCAGAGAGGACAGGACATGGTCTC	4052
Oy	3892	CTTATAGGAAGTTAACTTAATAGTGGCTTTATGTGCTTCAGAGAAGCAATTTCTCTGAGC	3951
Dp	4053	CTTATAGGAAGTAACCTTCTCAGTGGCTTTATGTGTGATCTTCAAGATTCATTTCTCTGAGT	4112
Oy	3952	GAGACAGTGAAGAAAGTGAATATGATCATGGAATACTCAACTCTTAATTAATTTCTGTAT	4011

Db	4113	GAGACCTGAAAGATGGAATATGATATACGGGAAATCAACTTTATTATTTAGATTCTGTGA	41722
Qy	4012	AATGAAACCAAGTTTGTGTATATATCTCTGTGTGAAACTTTAAAGTTTCAATACC	40711
Db	4173	AATGAATCCAGTTTGTGTTAATATCCACTGTGTAGAGACATACAAAGTTTCAAAATTT	42322
Qy	4072	CAAGATGCTTCAGTGTCCATATAGAGATTACTATGAGCCAAAGAGACAGCGCGTGAGATG	41311
Db	4233	CGAATGTTCAGTATCTGTATATGAGACTTACTATGAAACCAAGAGACAGGCACTGTCCAGAC	42322
Qy	4132	TACAACTCTGAAGTGAAGCTGTCTCTCTGTGACCTTTGACAGTAGTAGTCCAGGCGTGCCT	41911
Db	4293	TACAAACCCAGGTGAAGCTGTCTCTCATGTATCTCT--CAGTCCAGACACCAACTGCAAG	43419
Qy	4192	CTTTGTGAGATGAGAGCTTCAGGCTCCCATCATCATCTTTCAGTCATTTTATTTTCTGT	42511
Db	4350	TCACACACGAGACGAGGACACGGAGCTCCCTTCGAGAGTCTTCGAGCTTCGTCTGTTC	44019
Qy	4252	TTCAAGCTTCGTATCTTATAGAACTTTGGCTGTGA	42817
Db	4410	TCGCTCTCTGTACTTTGTGCAACATTTGATATGA	44415

[illegible]

Quackenbush, J., Schriml, L.M., Staab, F., Suzuki, R., Tomita, M., Wagner, J., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H., Baldarelli, R., Barsh, G., Blake, J., Boffelli, D., Bojunga, N., Carninci, P., de Bonaldo, M.F., Brownstein, M.J., Bult, C., Fletcher, C., Fujita, M., Gariboldi, M., Gietz, S., Hill, D., Hofmann, M., Hume, D.A., Kamiya, M., Lee, N.H., Lyons, P., Marchionni, J., Mashima, J., Mazzarelli, J., Mombaur, P., Nordone, P., Ring, B., Ringwald, J., Rodriguez, I., Sakamoto, N., Sasaki, H., Sato, K., Schonbach, C., Seta, T., Shibata, Y., Storch, K.F., Suzuki, H., Toyooka, K., Wang, K.H., Weitz, C., Whitaker, C., Wilming, L., Wynshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawaji, H., Kohsaki, S. and Hayashizaki, Y.

Functional annotation of a full-length mouse cDNA collection  
Nature 409 (6821), 685-690 (2001)

11217851

JOURNAL MEDLINE PUBMED

TITLE

AUTHORS

5

The PANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.

Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs  
Nature 420, 563-573 (2002)

6 (bases 1 to 4725)

Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoaka, T., Hirozawa, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Kato, S., Kawai, J., Kojima, Y., Kondo, S., Kono, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N., Okazaki, Y., Saito, R., Saitoh, K., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y.

Direct Submission

Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gs.c.riken.go.jp, URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)

CNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.

Please visit our web site for further details.  
URL: http://phantom.gsc.riken.go.jp/  
Location/Qualifiers

1. 4725

/organism="Mus musculus"

/mol\_type="mRNA"

/strain="C57BL/6J"

/db\_xref="PANTOM DB:9930012E15"

/db\_xref="taxon:10090"

/clone="9930012E15"

/sex="female"

/tissue\_type="vagina"

/clone\_lib="RIKEN full-length enriched mouse cDNA library"

/dev\_stage="adult"

97..4423

misc\_feature

/note="putative  
weakly similar to alpha-2-macroglobulin (fragments)  
[Limulus polyphemus] (PIR|A36260, evidence: FASTV,  
51.&ID, 76.7&length, match=168) "

4707..4712

polya\_signal

/note="putative"

4725

polya\_site

/note="putative"

BASE COUNT 1334 a 1016 c 1066 g 1309 t

ORIGIN

Query Match 62.3%; Score 2668.8; DB 11; Length 4725;  
Best Local Similarity 77.7%; Pred. No. 0;  
Matches 3369; Conservative 0; Mismatches 877; Indels 94; Gaps 10;

1 ATGCAGGGCCACCGCTCTGACCGCCGCCACCTCTCTGCTGACCGCCGGCTG 60  
97 ATGCCAGTCGAGAGGCTCTGAGCCGCCACCTTCTCTGCTTGGCGGCTCGCTG 156  
61 GCCGTGCTCCCGGCTCTGCTTCTGCTGACGCCCCAGGATATCATCGCCCGAGCA 120  
157 GCC--GCCCGCGGGTCCGCTTCTGCTGACGCCCCAGGATATCATCGCCGAGCA 213  
121 AATGTACATTTGGGGTGGAGCTCTGGAACCTGCCCTTCA CAGGTGCTGGAAGGGG 180  
214 AATGTACATTTGGGGTGGAGCTCTGGAATAATACCCGCCACAGTCTTTAAAGCT 273  
181 GAGCTGTCAGACAGCATCAAACTCACTGCTCTGCTGGAAGCAGAAAGAGCTTT 240  
274 CAGGTGTCAGATAGCTTCCAAACAAATCGAGATCATCTGGAAGCAGAAAGAGCTTT 333  
241 GAAAAAGCTCTTTTAAGCACTTACTTCCATCACTACTTGAACAGTGCAGATAG 300  
334 CACAGAGGCATTTCAAGACTCTCGTTCCTCGGACACTTACCTGAGCAGTGCAGATAG 393  
301 ATTATGAGCTACGCTGTAACCGGACGTAACCGAGATGATTTTCTTAAATAGTACC 360  
394 ATTATGAGCTACATTAACGAGACATGAGATGAGATGATGATTTCTTCAACGAGCA 453  
361 CCTTATCATTTGAGACCAAGAAATATCTGCTTCAATTCATTCAAACAGAGCTTTTAC 420  
454 CGTTTAACTTTGAGAGCAAGATATCTGCTTCAATTCATTCAGACAGAGCATTTTAC 513  
421 AAGCCAAAGCAAGAGTGAAGTTGCAATTTGCACTTCTCTGATTTTAAAGCTTAC 480  
514 AAGCCAAAGCAAGAGTGAAGTTGCGGTGCTCACTCTGCTGATTTCAAGCTTAT 573  
481 AAACCTCTTTAAATTTCTTAAAGAACCCCAATCAAAATTTGATTCGAAGTGTG 540  
574 AAGACATCTGAGCAATTTTCAATTAAGATCGGAAGTCAAGTATCCAAAGTGTG 633  
541 TCACAAAGAGTGAAGTCTGAGTCAATTTCCAAACCTTTGAGCTATTTCCATCAATA 600  
634 TCTCAGAAAGGATCTTGAAGTCTTTCCAAACCTTTTCAATCTCTTAATCAATA 693  
601 CTGTGATCTGCTTATTCAGTTCAAGTGAATGACAGACATATTAATCAATTCATG 660  
694 TTTGGTGAAGTCTTCAATTCAGTTCAAGTGAATGACAGACATATTAATCAATTCATG 753  
661 GTTTCAGAAATGATTTTACCAAAATTTGAAGTGAATTTGACAGACCATTAATTTGCT 720  
754 GTCTAGAAATGATTTTACCAAAATTTGAAGTGAATTTGACAGACCATTAATTTGCT 813  
721 ATGAATTTCAAGCATTTAAATGATGATCAATCGGCAAGTATATATGGAAGCAGTG 780  
814 TTGAAGTCTTAAACATTTAAATGATGATGATGATGATGATGATGATGATGATGAT 873  
781 AAAGG--AGACGTAAAGCTTACATTTTACCTTTT--GGGGAAGAGAAAT 837  
874 GAAAGGAGGTCTGACATTAATTTTACCTTTTGGGGAAGAGAAAT 933  
838 ATTACAAAACATTTAAATGATGATGATGATGATGATGATGATGATGATGATGATG 897  
934 ATCAACAAAAGTTTGAATTAATGATGATGATGATGATGATGATGATGATGATGATG 993  
898 AAAATGTAAT-----GATTTCTTAAATGATGATGATGATGATGATGATGATG 939  
994 AAAAGGTAATTAATTTGAAACCGGTAAAGATGATGATGATGATGATGATGATGATG 1053  
940 GATCTATCTTCCCTGACAGATGAAATTTTAAACAGATGACAGATGATGATGATGATG 999  
1054 GACCGTCTTCTTCTGAGCCGCGACAGATTTATGATGATGATGATGATGATGATGATG 1113

1000 ATTTCAGAAAATGTAAGCACTAATGTGTTCTTCAAGCAACATGATTTACATCATTTGAGTTT 1059  
1114 ATCTCAAAATGCGAGACCAATGTGTTTTCAGAGCAACATGATTTACATCATTTGAGTTT 1173  
1060 TTGTATTAATCTACTGTCTTGAAGCAATCTCTCAACTTTCACAGCCACTGTGAAGTTACT 1119  
1174 TTGTATTAATCTACTGTCTTGAAGCAATCTCTCAACTTTCACAGCCACTGTGAAGTTACT 1233  
1120 GGTGTGATGCGCAACCACTGCTCTTGAAGCAACCAATGATTTGATGATTAACG-T 1178  
1234 GGTGTGATGCGCAACCACTGCTCTTGAAGCAACCAATGATTTGATGATTAACG-T 1293  
1179 GACAAGAGAAATCTAATCTGAGTACTGAGAGGCAATCTAAGCACTGAGAAATGGA 1238  
1294 GACCAGAGGAGAAATACCAACCCAGAG-----TCAGAGAGATCAGAGAAATGGA 1344  
1239 AGCTGTTCAAGAAATTAATTAATCTGTCCTTCCCAAGTGAATCTTTAAGATTGAATCCC 1298  
1345 TTATATCCAGACCCGTAATTAATCTATCCCAAAATGGAATCAATTAAGATTGAATCCC 1404  
1299 AATCTGAGAGATTCAGAGTACAGTACAGTGAAGGCCATTTCTGTGTAGTAAAGTAG 1358  
1405 AGTCAATGCCATTTCCGCGAGCTACCACTGAAGGCTTAATTTCTGTGTAGTAAAGTAG 1464  
1359 CATGCACTTCAATAGTCTGTTTAAGTCTCTAGTGAACATACATCCAACTAAAAACAG 1418  
1465 TGTGACATCCAGCATATGTTCACTCTCTAGTGAACATACATCCAACTAAAAACAG 1524  
1419 AGATGAAAAATTAAGTGGGATGCGCTTTGAGTGTGTGTAGTGAACAAACGAT 1478  
1525 AGATGAAAAATTAAGTGGGATGCGCTTTGAGTGTGTGTAGTGAACAAACGAT 1584  
1479 GAAGAGTGAAGTATAGTATAGTATCCAGGGAACAGTGTGTGTAGTGAACAAACGAT 1538  
1585 CAAGAGTGAAGTATAGTATAGTATCCAGGGAACAGTGTGTGTAGTGAACAAACGAT 1644  
1539 TTCAACAATGTTCTCTTAAACACAGAAATTTCTGAGTCCAAAGGCTGTGTATTTGT 1598  
1645 TTCAACAATGTTCTCTTAAACACAGAAATTTCTGAGTCCAAAGGCTGTGTATTTGT 1704  
1599 GTATTAATTAAGAGATGAGGAAATTAATGAATGTTCTTAAATTTCTGTGTACCT 1658  
1705 ATATATATATGAGAGATGAGGAAATTAATGAATGTTCTTAAATTTCTGTGTACCT 1764  
1659 TGTATTAATTAAGAGATGAGGAAATTAATGAATGTTCTTAAATTTCTGTGTACCT 1718  
1765 TGTATTAATTAAGAGATGAGGAAATTAATGAATGTTCTTAAATTTCTGTGTACCT 1824  
1719 AGTCTCTTGAAGTCTCTGTAACACAGCTGATCCATGATGAGGATTTGTAGTGTGA 1778  
1825 GGTCTCTTGAAGTCTCTGTAACACAGCTGATCCATGATGAGGATTTGTAGTGTGA 1884  
1779 CAAGAAGTGAATCTGATGAATGCTCTTAATGATTAATCAATGAAGAAATGTGTCCATGA 1838  
1885 CAAGAAGTGAATCTGATGAATGCTCTTAATGATTAATCAATGAAGAAATGTGTCCATGA 1944  
1839 GTTGAAGTGAATCTGATGAATGCTCTTAATGATTAATCAATGAAGAAATGTGTCCATGA 1898  
1945 ACTGAAGTGAATCTGATGAATGCTCTTAATGATTAATCAATGAAGAAATGTGTCCATGA 2004  
1899 TCAGAGATGAGGATCTGAGGATTTGAGAGATGCAAACTCGAAGATTAATGAAGG 1958  
2005 TCAGAGATGAGGATCTGAGGATTTGAGAGATGCAAACTCGAAGATTAATGAAGG 2064  
1959 TGTATTAATTAAGAGATGAGGATTTGAGAGATGCAAACTCGAAGATTAATGAAGG 2018  
2065 AGTCAAGATTAAGAGATGAGGATTTGAGAGATGCAAACTCGAAGATTAATGAAGG 2124  
2019 AGATTAATTAAGAGATGAGGATTTGAGAGATGCAAACTCGAAGATTAATGAAGG 2078  
2125 TGTATTAATTAAGAGATGAGGATTTGAGAGATGCAAACTCGAAGATTAATGAAGG 2184

2079 TTGATTTGGCTAGACACCAATGAGTTACAGAGTTTACCAAGATTTGAGTACTGT 2138  
2185 ATGATTTGGCTAGACACCAATGAGTTACAGAGTTTACCAAGATTTGAGTACTGT 2244  
2139 ACCGATTTCTATCACTTCTTGGGTGCTACTGTTTGTGATATCTGAGAGACCTGGGCT 2198  
2245 ACCGATTTCTATCACTTCTTGGGTGCTACTGTTTGTGATATCTGAGAGACCTGGGCT 2304  
2199 TGACATTAACAATCTACTGAGAGTCCAAAGCTTCCAAACATTTTTCATTTTGTAA 2258  
2305 TGGGTTAACAACCTTCCAGAGAGTCCAAAGCTTCCAAACATTTTTCATTTTGTAA 2364  
2259 TCTTCCCTACTCTGTTATCCAGAGTGAAGATTTGTTGAAATTAATTAATTAATTA 2318  
2365 TCTTCCCTACTCTGTTATCCAGAGTGAAGATTTGTTGAAATTAATTAATTAATTA 2424  
2219 TTGAAAGATGCGACATGAGTTAAGTAAATTAATTAATTAATTAATTAATTAATTA 2378  
2425 TTGAAAGATGCGACATGAGTTAAGTAAATTAATTAATTAATTAATTAATTAATTA 2484  
2379 AATGACTTCAAGTGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2438  
2485 AATGACTTCAAGTGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2544  
2439 GATGAGGCAATCTGTTTCTTCCATCAGGCAACATCTGGAGAAATTTCTATCAC 2498  
2545 TAATGGGCTACTGTTTCTTCCATCAGGCAACATCTGGAGAAATTTCTATCAC 2604  
2499 AGTCAAGCTCTTCAACCTGCTTGTATGCTGTACCCCAATGATTTTATGTAAGGC 2558  
2605 AGTCAAGCTCTTCAACCTGCTTGTATGCTGTACCCCAATGATTTTATGTAAGGC 2664  
2559 TGAAGAAATGAAGAAATCAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2618  
2665 TGAAGAAATGAAGAAATCAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2724  
2619 ACAAGATACCTGAAACCTTGAATTTCTGATTTCTCTTAATTAATTAATTAATTA 2678  
2725 AGAAGATACCTGAAACCTTGAATTTCTGATTTCTCTTAATTAATTAATTAATTA 2784  
2679 AAGAGTCAATCACTGCAATTTGAAGATTTCTTGGTCTTCCATCAATGCTTGAAGCT 2738  
2785 AAGAGTCAATCACTGCAATTTGAAGATTTCTTGGTCTTCCATCAATGCTTGAAGCT 2844  
2739 ATTGATCGAGTCTTATGAGTGTGAGTGAACAGAAATTAATTTCTCCAAATAT 2798  
2845 ACTGATCGAGTCTTATGAGTGTGAGTGAACAGAAATTAATTTCTCCAAATAT 2904  
2799 TTACATTTTGAATTAATGATTAATTAATTAATTAATTAATTAATTAATTAATTA 2858  
2905 TTACATTTTGAATTAATGATTAATTAATTAATTAATTAATTAATTAATTAATTA 2964  
2859 TCTTCAATTAATGAGGCAAGTTAATTAATTAATTAATTAATTAATTAATTAATTA 2918  
2965 TCTTCAATTAATGAGGCAAGTTAATTAATTAATTAATTAATTAATTAATTAATTA 3024  
2919 TTTCAAGTCTTGGGAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2978  
3025 TTTCAAGTCTTGGGAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3084  
2979 AAGATGTTCTTGAAGCCGATCTTAATTAATTAATTAATTAATTAATTAATTAATTA 3038  
3085 AAGATGTTCTTGAAGCCGATCTTAATTAATTAATTAATTAATTAATTAATTAATTA 3144  
3145 AATGATGTTCTTGAAGCCGATCTTAATTAATTAATTAATTAATTAATTAATTAATTA 3204  
3099 GATTCATGAGGATTAAGTGAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3158  
3205 GATTCATGAGGATTAAGTGAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3264  
3159 AATTCCTCTCTGAGATTAAGAAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3218

```

Db      3265 GACTCTGCTCGGAGTACAAAAGTATACGCTAATATGATGACAAAGCTTCATCA 3324
Oy      3219 TTTTGGAGTCTGAATTCAGTAAATTTGAGCAATTAATCTGACCCCTTAAC 3278
Db      3325 GTTTTGGATTGAAATTCAGCAGAGAAATTCGACAAATTAATCTTGCAATTAATC 3384
Oy      3279 TTAATGATTCATCATAGTGGGAGTCTTAAGCGAAGAACTTGAATATGCTGA 3338
Db      3385 CATATGCTCTGTCACAGTCGGGAGCCCTTAAGCAGAGAGGCTTTGAATCTGTGATGA 3444
Oy      3339 GAGAGAGAACAAAGAAAGTGGCATTAATTTGGGTGATCATAGAGTCCAACTTTCTGA 3398
Db      3445 GGGATCAGAAAAGAGAGAGACACAGATTCGTGTGATCAGAGACCTGCGCTCTCTGG 3504
Oy      3399 CTCCTGGCAGCAGCGCTCCCTGGATATTAAGTGGAGCCTAATGCACTGCTCAACATT 3458
Db      3505 GTCTTGAGCAGCCGCGCTCCGTGACATCGAAATCGCTTATGACCTGTGCGCACAC 3564
Oy      3459 CTTAACAATTCAGACTTCTGAGGAAATCCCAATTAAGAGTGGCTTAAGCAGGCAAGAA 3518
Db      3565 TCTGCA---CCATGATCTGAGGAAATCCGGTTATGAGTGGCTCATCCAGCAAGAA 3621
Oy      3519 TAGCTTGGGTGTTTGCATCTACTAGAGATACACTGTGGCTTTAAGGCTCTGTCTGA 3578
Db      3622 CAGCCGGAGAGGTTTGTATCCACTCAGGATACTGTTGTGCGCTGAAAGCATTAATCTGA 3681
Oy      3579 ATTTGAGCCCTTAATGAATACAAAGAAAGCAATATCCAAAGTACCGGTACCGGCGCTAG 3638
Db      3682 ATTTTACGCCCTTGTCACAGAAAGAAATACAGATATCAATGACCGTCAAGGCGCCG 3741
Oy      3639 CTCACCAAGTCTT----- 3651
Db      3742 CATCCCAATCATACATTCGCAATGATGATCCGAACTCTTCTCTCACAGAA 3801
Oy      3652 ----CTTGCTGTGTGATACAGCCAAATGAGTAAATTTCCGAAATGTTTGGATTTC 3707
Db      3802 AGAGCTTACGCGCTGATCCATCCATCCGTTAATGTTTCTGCGCATGCGTCCGGAATTC 3861
Oy      3708 TATTTGTGCTCAATGTTGATATTAATGAAGGCTTCTGGCTCTTCAAGAGCAAG 3767
Db      3862 TATTTGCCAGTTAATGTTGACTATTAACGTGAAGGTTTCTTCAAAAGACGAG 3921
Oy      3768 ATCTATCCAAATTCAGAAAGCTTTGATTTAGTGTGCTGTAAGAAATTAAGATGA 3827
Db      3922 ATCTACCGAAATTCAGAAAGTTTGAATTAAGCTGATTTGA---ATAATGAGAGCA 3978
Oy      3828 TCTCATCATGTGATTAATGATGTGTGACAAAGCTTTGCGGCCCGGTAGAGAGTGCAT 3887
Db      3979 CATATGATCACTGAATCTGAATGTGTGACAAAGTCACTTGGGTTCAAGAGACAGGCAT 4038
Oy      3888 GGCTCTTAATGAAGTTAACTATTAATGAGCTTATGAGCTTCAAGAGCAATTTCT 3947
Db      4039 GGTCTTAATGAAGTAACTTCTCAGTGTGATTAATGAGATCTTCAAGATTAATTCCT 4098
Oy      3948 GAGCGAGACAGTGAAGAAAGTGAATATGATATGAGAAATCAACTCTAATTTAGATTC 4007
Db      4099 GAGTGAAGACCTGAAAGAAAGTGAATATTAACGGAATCACTCACTTATTTAGATTC 4158
Oy      4008 TGTAAATGAATCCAGTTTGTGTTAATAATCTGCTGTGAGAACTTTAAAGTTTCAA 4067
Db      4159 TGTAAATGAATCCAGTTTGTGTTAATAATCTGCTGTGAGAACTTTAAAGTTTCAA 4218
Oy      4068 TACCAAGATGCTTCAAGTGTCAATGATGATTAATTAAGCAAGAGACAGCGGTGAG 4127
Db      4219 TATTTGAGATGTTGATCTGTATATGATTAATTAAGCAAGAGACAGCGAGTGG 4278
Oy      4128 AAGTTCAACTCTGAAGTGAAGCTCTCTCTGTGACTTTGAGAGTGAAGTCAAGGCTG 4187
Db      4279 AAGCTCAACACCCAGGTGAAGTGTCTCTGTGATCTT---CAGTCCAGACCAACTG 4335
Oy      4188 CCGTCTGTGAGATGAGCTTCAAGCTTCCATCATCTTCAAGTATTTTATTTT 4247

```

```

Db      4336 CAATGACACACAGCAGGACCGACGACTCCCTTGACGTTCTTGAGCTTCTGCTT 4395
Oy      4248 CTGTTCAAGCTTCTGATCTTATTAATGAACCTTGGCTGTA 4287
Db      4336 TTGCTCCCTCTTCTGATCTTTGTCACAAATGATTAATGA 4435

RESULT 3
AL554092
LOCUS
DEFINITION
AL554092 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
clone CS0D1081YB18 5-PRIME, mRNA sequence.
ACCESSION
AL554092
VERSION
AL554092.2 GI:11275905
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1 (bases 1 to 1201)
Li, W.B., Gruber, C., Jessup, J., and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished
JOURNAL
On Feb 15, 2001 this sequence version replaced gi:12894544.
COMMENT
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 Evry cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 3578.f for
more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CS0D1081YB18&cluster=3578.f. Contact :
Feng Liang Email : fliang@life.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Faraday Avenue Genoscope sequence ID : CS0D1081YB18.

FEATURES
source
1..1201
location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0D1081YB18"
/issue_type="PLACENTA COT 25-NORMALIZED"
/clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
/note="First strand cDNA was primed with a NotI-oligo (dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."
BASE COUNT
359 a 238 c 236 g 293 t 75 others
ORIGIN

Query Match 20.8%; Score 893; DB 9; Length 1201;
Best Local Similarity 93.9%; Pred. No. 2.3e-187;
Matches 947; Conservative 39; Mismatches 14; Indels 9; Gaps 6;

Oy      1 ATGCAAGGCGCAACGCTCTGACCGCGCAACCTCTCTGCTGTGACCGCGCGTG 60
Db      92 ATGCAAGGCGCAACGCTCTGACCGCGCAACCTCTCTGCTGTGACCGCGCGTG 151
Oy      61 GCCGTGCTCCCGGCTCGGTTTCTGCTGACAGCCCGAGGATCATCAGGCCGAGAGA 120
Db      152 GCCGTGCTCCCGGCTCGGTTTCTGCTGACAGCCCGAGGATCATCAGGCCGAGAGA 211
Oy      121 AATGTGACTATTTGGGTGAGACTTCTGGAACAATGCTTCAAGGTGATCTGGAAGCG 180
Db      212 AATGTGACTATTTGGGTGAGACTTCTGGAAMAATGCTTCAAGGTGATCTGGAAGCG 271
Oy      181 GAGGTGCTCAAGACGATCAACCTCATCTGTCTGCTGGAAGAGAGAGTCTT 240
Db      272 GAGGTGCTCAAGACGATCAAAATTCATCTGTCTGTGCTGGAAGAGAGTCTT 331
Oy      241 GAAAAAGGCTTTTAAGACACTTCTTCATCACTCACTGAAACAGTGCAGATGAG 300

```



Db 332 GAAAAAGCTCTTTTAAAMACTTAMTCTTCCATWATMCTGAAACATGACATGAG 391  
 Qy 301 ATTATGAGTACGTAACGGAGCTACCCAGATGAGATTTATTTCTTATAGTACC 360  
 Db 392 ATTATGAGTACGTAACGGAGCTACCCAGATGAGATTTATTTCTTATAGTACC 451  
 Qy 361 CGTTATGATTTGAGACCAAGATATCTGTTCTTCAATTCACAGACAGAGCTTTATC 420  
 Db 452 CGTTATGATTTGAGACCAAGATATCTGTTCTTCAATTCACAGAGAGCTTTATC 511  
 Qy 421 AAGCCAAAGCAAGAGTGAATGTTGCGATTTGTACACTCTTCTCAGATTTTAACTTAC 480  
 Db 512 AAGCAAAAGCAAGAGTGAATGTTGCGATTTGTACACTCTTCTCAGATTTTAACTTAC 571  
 Qy 481 AAAACCTTTTAACTTCTCAGATTTGAGACCAAGATTTGATGATGAGAGTGGTGG 540  
 Db 572 AAAACCTTTTAACTTCTCAGATTTGAGACCAAGATTTGATGATGAGAGTGGTGG 631  
 Qy 541 TCACAAAGAGTATCTTGAAGTCAATTCACAACTTTTCACTATCTTCCATCAATA 600  
 Db 632 TCACAAAGAGTATCTTGAAGTCAATTCACAACTTTTCACTATCTTCCATCAATA 691  
 Qy 601 CTGTGATCTGTCTATTCAGATTTCAAGTGAATGACGACATTTATTCATCATTTTAC 660  
 Db 692 CTGTGATCTGTCTATTCAGATTTCAAGTGAATGACGACATTTATTCATCATTTTAC 751  
 Qy 661 GTTTCAGATTTGATTTCAAGATTTGAGTGAATGACGACATTTATTCATCATTTTAC 720  
 Db 752 GTTTCAGATTTGATTTCAAGATTTGAGTGAATGACGACATTTATTCATCATTTTAC 811  
 Qy 721 ATGATTTTCAAGATTTTAACTTACATTCACGCAAGATTTATGATGAGAGAGTGG 780  
 Db 812 ATGATTTTCAAGATTTTAACTTACATTCACGCAAGATTTATGATGAGAGAGTGG 870  
 Qy 781 AAAGAGAGCTGACGCTTACATTTTACCTTTATCTTTGGGAAAGAAAGAAATTT 840  
 Db 871 AAAGAGAGCTGACGCTTACATTTTACCTTTATCTTTGGGAAAGAAAGAAATTT 930  
 Qy 841 ACAAACATTTTAACTTAACTTGAATGATGCAAACTCTTTTAACTTGAATGAGAGTGA 900  
 Db 931 ACAAACATTTTAACTTAACTTGAATGATGCAAACTCTTTTAACTTGAATGAGAGTGA 990  
 Qy 901 AATGATGATTTTCAAGATTTTAACTTGAATGATGCAAACTCTTTTCAAGATTTTCAAG 960  
 Db 991 AATGATGATTTTCAAGATTTTAACTTGAATGATGCAAACTCTTTTCAAGATTTTCAAG 1045  
 Qy 961 GTAGAAATTTTCAAGATTTTAACTTGAATGATGCAAACTCTTTTCAAGATTTTCAAG 1009  
 Db 1046 GTAGAAATTTTCAAGATTTTAACTTGAATGATGCAAACTCTTTTCAAGATTTTCAAG 1091  
 RESULT 4  
 BX391819/c 941 bp mRNA linear EST 13-MAY-2003  
 LOCUS BX391819 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA  
 DEFINITION clone CS001081YB18 3-PRIME, mRNA sequence.  
 ACCESSION BX391819  
 VERSION BX391819.1 GI:30627404  
 KEYWORDS EST.  
 ORGANISM Homo sapiens (human)  
 SOURCE Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 REFERENCE 1 (bases 1 to 941)  
 AUTHORS Li, W.B., Gruber, C., Jessee, J. and Polayes, D.  
 TITLE Full-length cDNA libraries and normalization  
 JOURNAL Unpublished  
 COMMENT Genoscope - Centre National de Sequencage  
 BP 191 91006 Evry cedex - France  
 Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr  
 Library was constructed by Life Technologies, a division of  
 Invitrogen. This sequence belongs to sequence cluster 3578.f For

more information about this cluster, see  
<http://www.genoscope.cns.fr/cgi-bin/cluster.cgi?seq=CS001081YB18>  
 Contact : Feng Liang Email : fliang@life.technet.com url :  
<http://fulllength.invitrogen.com/InvitrogenCorporation1600>  
 Faraday Avenue Genoscope sequence ID : CS001081YB18\_C502740\_1.

FEATURES  
 source  
 1. 941  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="CS001081YB18"  
 /issue\_type="PLACENTA COT 25-NORMALIZED"  
 /note="1st strand cDNA was primed with a NotI-oligo (dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."

BASE COUNT 271 a 214 c 181 g 266 t 9 others

ORIGIN  
 Query Match 20.4% Score 872.8; DB 13; Length 941;  
 Best Local Similarity 96.7%; Pred. No. 6.8e-183;  
 Matches 909; Conservative 0; Mismatches 28; Indels 3; Gaps 2;

Db 2820 TAAAAAGAAACAAGTACAGATTAATTTGAAAGAAAGCTCTTCAATTTATGAGGCAAG 2879  
 Qy 939 TAAAGAAACACTGACAGATATTGAAAGAAAGAGCTCTTCAATTTATGAGGCAAG 880  
 Db 2880 TTAACAGAGAACTTCTCTATCAGAGAGAGAGTCTCTTCAAGTCTTTGGAAATTA 2939  
 Qy 879 GTACAGAGAGAGCTTCTCTATCAGAGAGAGAGTCTCTTCAAGTCTTTGGAAATTA 821  
 Db 2940 TGAACCTTCTGAGAGAGCTTCTCTATCAGAGAGAGAGTCTCTTCAAGTCTTTGGAAATTA 2999  
 Qy 820 TGAACCTTCTGAGAGAGCTTCTCTATCAGAGAGAGAGTCTCTTCAAGTCTTTGGAAATTA 761  
 Db 3000 TCCCTTCAATGATTAATGATCAGATGATGTTTACAGAGAGAGAGTCTTCAAGTCTTTAAAGACA 3059  
 Qy 760 TCCCTTCAATGATTAATGATCAGATGATGTTTACAGAGAGAGAGTCTTCAAGTCTTTAAAGACA 701  
 Db 3060 TCAAGAAATCAACGAGTGAATTTGGATCCAGAGAGAGAGTCTTCAAGTCTTTAAAGACA 3119  
 Qy 700 TCAAGAAATCAACGAGTGAATTTGGATCCAGAGAGAGAGTCTTCAAGTCTTTAAAGACA 641  
 Db 3120 TGGCAATTAAGTCCAGTACACTTCAAGTCTTCAAGTCTTTCAAGTCTTTAAAGACA 3179  
 Qy 640 TGGCAATTAAGTCCAGTACACTTCAAGTCTTCAAGTCTTTCAAGTCTTTAAAGACA 581  
 Db 3180 AAAGTATCAGCTTCAAGTATGATGCAAGAGTCTTCAAGTCTTTCAAGTCTTTAAAGACA 3239  
 Qy 580 AAAGTATCAGCTTCAAGTATGATGCAAGAGTCTTCAAGTCTTTCAAGTCTTTAAAGACA 521  
 Db 3240 TAAGAGATTTCAAGCAATTAATCTAGCTTCAAGTCTTTCAAGTCTTTAAAGACA 3299  
 Qy 520 TAAGAGATTTCAAGCAATTAATCTAGCTTCAAGTCTTTCAAGTCTTTAAAGACA 461  
 Db 3300 GAGTCTTCAAGCAATTAATCTAGCTTCAAGTCTTTCAAGTCTTTAAAGACA 3359  
 Qy 460 GAGTCTTCAAGCAATTAATCTAGCTTCAAGTCTTTCAAGTCTTTAAAGACA 401  
 Db 3360 CATGCAATTTGGAGTCTTCAAGTCTTCAAGTCTTTCAAGTCTTTAAAGACA 3419  
 Qy 400 CATGCAATTTGGAGTCTTCAAGTCTTCAAGTCTTTCAAGTCTTTAAAGACA 341  
 Db 3420 GATATTTGAAGTTCAGAGCTTCAAGTCTTCAAGTCTTTCAAGTCTTTAAAGACA 3479  
 Qy 340 GATATTTGAAGTTCAGAGCTTCAAGTCTTCAAGTCTTTCAAGTCTTTAAAGACA 281  
 Db 3480 GGGATTCCTCAATTTATGAGTCTTCAAGGCAAGAAAGAGTCTTGGAGTCTTTGATC 3539  
 Qy 280 GGGATTCCTCAATTTATGAGTCTTCAAGGCAAGAAAGAGTCTTGGAGTCTTTGATC 221



QY 3540 TACTGAGTATACCACTGCTGCTTTAAAGGCTCTGCTGTAATTTGACGCCCTATGAAATC 3599  
| | | | |  
DB 220 TACTGAGTATACCACTGCTGCTTTAAAGGCTCTGCTGTAATTTGACGCCCTATGAAATC 161  
| | | | |  
QY 3600 AGAAGAGCAAAATATATCAAGTACCGGTGACGGGGCTTACTACCAAGTCCCTGCTCT 3659  
| | | | |  
DB 160 AGAAGAGCAAAATATATCAAGTACCGGTGACGGGGCTTACTACCAAGTCCCTGCTCT 101  
| | | | |  
QY 3660 GGATACGCCCAATGAGCTTATATATTTCCGCAATGGTTTGGATTTGCTATTTGTCAGCT 3719  
| | | | |  
DB 100 GGATACGCCCAATGAGCTTATATATTTCCGCAATGGTTTGGATTTGCTATTTGTCAGCT 41  
| | | | |  
QY 3720 CAATG--TTGTATATATATGTAAGGCTTCTGGCTCTTCTA 3757  
| | | | |  
DB 40 CAATGTTTGTATATATATGTAAGGCTTCTGGCTCTTCTA 1  
| | | | |  
RESULT 5  
BUS17328 879 bp mRNA linear EST 12-SEP-2002  
LOCUS AGENCOURT.10164521 NIH\_MGC\_71 Homo sapiens cDNA clone IMAGE:6514905  
DEFINITION 5', mRNA sequence.  
ACCESSION BUS17328  
VERSION BUS17328.1 GI:22824854  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
1 (bases 1 to 879)  
NIH-MGC http://mgi.nci.nih.gov/  
National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished  
Contact: Robert Strausberg, Ph.D.  
Email: cge@bbs-rcmail.nih.gov  
Tissue Procurement: ATCC  
CDNA Library Preparation: Life Technologies, Inc.  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
http://image.llnl.gov  
Plate: LLM14090 row: P column: 10  
High quality sequence stop: 663.  
Location/Qualifiers  
1. 879  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:6514905"  
/tissue\_type="leiomyosarcoma"  
/lab\_host="DH10B (phage-resistant)"  
/clone\_lib="NIH\_MGC\_71"  
/note="Organ: uterus; Vector: pCMV-Sport6; site: 1; NotI;  
site 2: SalI; Cloned unidirectionally. Primer: Oligo dT.  
Average insert size 2.1 kb." 1 others  
BASE COUNT 248 a 171 c 206 g 253 t  
ORIGIN  
Query Match 16.6%; Score 712.4; DB 13; Length 879;  
Best Local Similarity 91.8%; Pred. No. 3e-147;  
Matches 807; Conservative 0; Mismatches 17; Indels 55; Gaps 3;  
QY 3207 AGAGCTATCCATTTTGGAGTCTGCAATTCAGTAGAGAAATTTTCAGCAATTAATCTT 3266  
| | | | |  
DB 1 AAGAGCTATCCATTTTGGAGTCTGCAATTCAGTAGAGAAATTTTCAGCAATTAATCTT 60  
| | | | |  
QY 3267 ACCCTTATACCTATGATGTCATGTCAGTGGGAGTCTCTAAAGCAAGAGCTTTGAA 3326  
| | | | |  
DB 61 ACCCTTATACCTATGATGTCATGTCAGTGGGAGTCTCTAAAGCAAGAGCTTTGAA 120  
| | | | |  
QY 3327 TATGCTGATGAGAGCAAGCAAGAGGTGCAATTCCTGGGTCTATCAGAGTC 3386  
| | | | |

DB 121 TATGCTGATGAGAGCAAGCAAGAGGTGCAATTCCTGGGTCTATCAGAGTC 180  
| | | | |  
QY 3387 CAACCTTTCTGACTCTCTGACAGCCAGCGCTCCCTGATATGAAAGTTGACGCTATGCACT 3446  
| | | | |  
DB 181 CAACCTTTCTGACTCTCTGACAGCCAGCGCTCCCTGATATGAAAGTTGACGCTATGCACT 240  
| | | | |  
QY 3447 GCTCTCACCTTCTTACAAATTTGACAGCTTCTGAGGGAATCCCAATATGAGGTGCTAAG 3506  
| | | | |  
DB 241 GCTCTCACCTTCTTACAAATTTGACAGCTTCTGAGGGAATCCCAATATGAGGTGCTAAG 300  
| | | | |  
QY 3507 CAGGAAAGAAATACCTTGGGTGTTTGGATCTCTCAGAGTATACCACTGCTGCTTTAA 3566  
| | | | |  
DB 301 CAGGAAAGAAATACCTTGGGTGTTTGGATCTCTCAGAGTATACCACTGCTGCTTTAA 360  
| | | | |  
QY 3567 GCTCTGCTGTAATTTGACGCCCTATATGATACAGAAAGCAATATACCAAGTACCGT 3626  
| | | | |  
DB 361 GCTCTGCTGTAATTTGACGCCCTATATGATACAGAAAGCAATATACCAAGTACCGT 420  
| | | | |  
QY 3627 GACGGGGGCTTACTGACCAAGTCTT----- 3651  
| | | | |  
DB 421 GACGGGGGCTTACTGACCAAGTCTTAAAGTTCTGATGACACACCAACCGCTTACT 480  
| | | | |  
QY 3652 -----CTTCTGCTGCTACAGCCATGCGATTAATTTCCGCAATG 3695  
| | | | |  
DB 481 CCTTCAGACAGCAGCTTCTGCTGCTACAGCCATGCGATTAATTTCCGCAATG 540  
| | | | |  
QY 3696 TTTGATTTGCTATTTGTCAGCTCAATGTTGATATATGAAAGGCTTCTGGCTTC 3755  
| | | | |  
DB 541 TTTGATTTGCTATTTGTCAGCTCAATGTTGATATATGAAAGGCTTCTGGCTTC 600  
| | | | |  
QY 3756 TAGAAGAGCAAGATCTATCCAAATCAAGAGCTTTGATTTGATGTTGCTGTAAGA 3815  
| | | | |  
DB 601 TAGAAGAGCAAGATCTATCCAAATCAAGAGCTTTGATTTGATGTTGCTGTAAGA 660  
| | | | |  
QY 3816 AATATAAGATGATCTCAATCATGTCGATTTGATGTCGATACAGCTTTTCCGCCCGG 3875  
| | | | |  
DB 661 AATATAAGATGATCTCAATCATGTCGATTTGATGTCGATACAGCTTTTCCGCCCGG 720  
| | | | |  
QY 3876 TAGAGTGGCAATGGCTCTTATGAGTAACTTAAGTGGCTTTATGTCCTTCA 3935  
| | | | |  
DB 721 TAGAGTGGCAATGGCTCTTATGAGTAACTTAAGTGGCTTTATGTCCTTCA 780  
| | | | |  
QY 3936 AGCAATTTCT--CTGAGCAGACAGTGAAGAAAGTGAATATGATCATGGAATCAACC 3994  
| | | | |  
DB 781 AGCAATTTCTCTGAGCAGACAGTGAAGAAAGTGAATATGATCATGGAATCAACC 840  
| | | | |  
QY 3995 TCT--ATTGATTTCTGTAATGAACCCAGTTTGTG 4030  
| | | | |  
DB 841 CCTTATTTNAGATTCTGTAATGAACCCAGTTTGTG 879  
| | | | |  
RESULT 6  
BX386457 978 bp mRNA linear EST 08-MAY-2003  
LOCUS BX386457 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA  
DEFINITION clone CS01081YB18 5-PRIME, mRNA sequence.  
ACCESSION BX386457  
VERSION BX386457.1 GI:30457430  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
1 (bases 1 to 978)  
Li, W.B., Gruber, C., Jessee, J. and Polayes, D.  
Full-length cDNA libraries and normalization  
Unpublished  
Contact: Genoscope  
Genoscope - Centre National de Sequencage  
BP 191 91006 EVRY cedex - France  
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr  
Library was constructed by Life Technologies, a division of  
Invitrogen. This sequence belongs to sequence cluster 3578.f for

more information about this cluster, see <http://www.genoscope.cns.fr/cgi-bin/cluster.cgi?seq=CSIA1021ZB100P1&cluster=3578.f>. Contact : Feng Liang Email : [liang@liferesearch.com](mailto:liang@liferesearch.com) URL : <http://fulllength.invtirgen.com> Invitrogen Corporation 1600 Faraday Avenue Genoscope sequence ID : CSIA1021ZB100P1 Location/Qualifiers

BASE COUNT ORIGIN	a	c	g	t	others
	374	151	179	220	54

Query Match	16.6%;	Score 710.8;	DB 13;	Length 978;
Best Local Similarity	84.6%;	Pred. No. 6.9e-147;		
Matches 758;	Conservative 46;	Mismatches 90;	Indels 2;	Gaps 2

Oy	66	GGCCCAACGGCTCCTGACCGCCGCCACCTCTCTGTCAGTGCACCGCGCGCTGGCCGTG
Db	143	GGGCCACGCGTCTCTGAACGAGACAAATCTCTGCGTGTGCACCGCGCGCTGGCCGTG
Oy	126	GCTCCCGGCGCTCGGTTTCTGTGTGACGAGCCCGAGGAAATCATCAGCGCCGAGAAATGTG
Db	203	GCTCCGCGGCGMTGGTGTWTGTGTGAGAGCCGAGGAAATTAAGGACCGGAGGAAATGTG
Oy	186	ACTATTTGGGGGTGGAGCTTCTGGAAACCTGCCCTTCAACAGTGTGACTGTGAAGCGGAGACTG
Db	262	ACATTTGGGGGTGGAGCTTCTGGAAACCTGCCCTTCAACAGTGTGACTGTGAAGCGGAG-SWG
Oy	246	CTCAGACAGCATCAAACTCTCATCTGTCTGTCTCTGAGCAGAGAGAGAGTCTTTGAAAAA
Db	322	CTCAGACAGCATCAAAATTTAAAGATCAGTATGAGAAAGAAAGATTTTGA AAAA
Oy	306	GGCTTTTAAAGCACTTACTCTTCCATCATCACTCTGAAACAGTGCAGATGAGATTAT
Db	382	AGATCTTTTAAARAAAAAATAACAAAAAAGGCAAGTGAAGAAATAT
Oy	366	GAGCTACGTGTAAACCGGACGTACCCAGATGATTTTATTTCTTAATAGTACC CGCTTA
Db	442	AAACMAAGTCTAAAGGAAGTACCCAGATGATATTTATCTCTTAAGAAACAGCGMTA
Oy	426	TCAATTTAGACCAAGAAATATCTGTTTCAATTCACACAGACAGAGCCTTATACAGCA
Db	502	AAAAATGAAAAACAAAAAATAAGAAATTAAGAAAAAABAAAGGATTTAAACAGCCA
Oy	486	AAGCAAGAAATGAAAGTTGCAATGTGTACACTCTTCTCAGATTTTAAAGCTTACAAAC
Db	562	AAGCAAGAAATGAAAGTTTGGCAATGTGTACACTCTTCTCAGATTTTAAAGCTTACAAAC
Oy	546	TCTTTAAACATTTCTCATTAAGACCCCAATCAAAATTTGATTCACAGTGTGTCAAA
Db	622	ACTTTAAACATTTTCAATTAAGAAACCAAAATTTAAATTTGATTCACAGTGTGTCAAA
Oy	606	CAAGTGAATCTTGAGTCATTTCCAAACTTTTGAGCTATCTTCCCATCCAAATCTTGCT
Db	682	CAAGTGAATCTTGAGTCATTTCCAAACTTTTGAGCTATCTTCCCATCCAAATCTTGCT
Oy	666	GACTGTCTAATTCAGTTCACAGTGAATGACAGACATATTTATCAATCAATTTCAAGTTTCA
Db	742	GACTGTCTAATTCAGTTCACAGTGAATGACAGACATATTTATCAATCAATTTCAAGTTTCA
Oy	726	GAAATATGTATTAACAAATTTGAAGTACTTTGACAGACACCATATATTTGTTCTATGAAT
Db	802	GAAATATGTATTAACAAATTTGAAGTACTTTGACAGACACCATATATTTGTTCTATGAAT

Oy	727	CTTAGACATTTAAATGGTACCATCGCAAGATATACATXGGAACCAGTGAAAAGA	786
Db	803	TCTNAGCATTAAATGGTACTATCATTTGCCAARSATATAGATTTSTAAGTAYGAAATCA	862
Oy	787	GACGTAAACGGTACATTTTTACCTTTATCCTTTTGGGG-AAAGAAAGAAAAATATTACANA	845
Db	863	GACTTAACCTKACATTTTKACCTTTTCCKKTGGGGSAAKAKAAAAATATTACANA	922
Oy	846	AACATTTAAGATAATGGATCTGCAGAACTTCTCTTTTAATGATGAAGAGATGANA	901
Db	923	AACATTTAAGATAATGSATCTGCAGAAATTTCTCTTTTAATGATGAAGAGATGANA	978

[illegible]

**REFERENCE**  
1 (bases 1 to 667)  
NCI-CCAP <http://www.ncbi.nlm.nih.gov/ncicgap>.  
**AUTHORS**  
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
**TITLE**  
Tumor Gene Index  
**JOURNAL**  
Unpublished  
**COMMENT**  
Contact: Robert Strausberg, Ph.D.  
Email: cgapsb-r@mail.nih.gov  
Tissue Procurement: Dr. Jose Mercuende  
cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa  
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa  
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa  
Clone Distribution: Clone distribution information can be found  
through the T.M.A.G.E. Consortium/MLNT at: <http://image.lnl.gov>  
Seq primer: M13 FORWARD  
POLYA=Yes.

BASE COUNT	194 a	113 c	139 g	221 t
ORIGIN				



Qy 4231 TCAGTCATTTTATTTCTGTTTCAAGCTTCTGATCTTATGAACTTTGCTGTGA 4287  
 Db 338 TCAGTCATTTTATTTCTGTTTCAAGCTTCTGATCTTATGAACTTTGCTGTGA 282  
 RESULT 9  
 BX386456/c 988 bp mRNA linear EST 08-MAY-2003  
 LOCUS BX386456 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA  
 DEFINITION clone CS001081YB18 3-PRIME, mRNA sequence.  
 ACCESSION BX386456  
 VERSION BX386456.1 GI:30453437  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 REFERENCE 1 (bases 1 to 988)  
 AUTHORS Li, W.B., Gruber, C., Jessee, J., and Polayes, D.  
 TITLE Full-length cDNA libraries and normalization  
 JOURNAL Unpublished  
 COMMENT Contact: Genoscope  
 Genoscope - Centre National de Sequencage  
 BP 191 91006 EVRY cedex - France  
 Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr  
 Library was constructed by Life Technologies, a division of  
 Invitrogen. This sequence belongs to sequence cluster 3578.f For  
 more information about this cluster, see  
 http://www.genoscope.cns.fr/  
 cgi-bin/cluster.cgi?seq=CS1A1021ZB10NP1&cluster=3578.f. Contact :  
 Feng liang Email : fliang@life.techn.com URL :  
 http://fulllength.invitrogen.com/ Invitrogen Corporation 1600  
 Faraday Avenue Genoscope sequence ID : CS1A1021ZB10NP1.  
 FEATURES  
 source location/Qualifiers  
 1..988  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="CS001081YB18"  
 /issue\_type="PLACENTA COT 25-NORMALIZED"  
 /note="1st strand cDNA was primed with a NotI-oligo (dt)  
 primer. Five prime end enriched, double-stranded cDNA was  
 digested with Not I and cloned into the Not I and EcoR V  
 sites of the PCMVSPORT 6 vector. Library was normalized."  
 BASE COUNT 384 a 133 c 164 g 249 t 58 others  
 ORIGIN  
 Query Match 14.2%; Score 608.6; DB 13; Length 988;  
 Best Local Similarity 90.6%; Pred. No. 3.6e-124;  
 Matches 642; Conservative 12; Mismatches 53; Indels 2; Gaps 2;  
 Qy 3580 TTTGACGCTTATGATGATGAGAAAGGACAAATATCCAGTGACCGGCGCTTACG 3639  
 Db 988 TTTGACGCTTATGATGATGAGAAAGGACAAATATCCAGTGACCGGCGCTTACG 929  
 Qy 3640 TCACCAATCTCTGCTGTGTGTAGAGCAATGAGTATATTTCCGCAATGCTTTT 3659  
 Db 928 TCACCAATCTCTGCTGTGTGTAGAGCAATGAGTATATTTCCGCAATGCTTTT 869  
 Qy 3700 GGATTTGCTATTTGTCAGCTCAATTTTATATATGAGAGCTTTGGGCTTTCTTGA 3759  
 Db 868 GGATTTGCTATTTGTCAGCTCAATTTTATATATGAGAGCTTTGGGCTTTCTTGA 809  
 Qy 3760 AGACGAGATCTATCCAAATTCAGAGACCTTTGATTTAGATGTTGCTGTAAAGAAAT 3819  
 Db 808 AGACGAGATCTATCCAAATTCAGAGACCTTTGATTTAGATGTTGCTGTAAAGAAAT 749  
 Qy 3820 AAAAGTATCTCAATCATGATGATTTGATGATGCTTACAGCTTTGGGCGCGGCTAGG 3879  
 Db 748 AAAAGTATCTCAATCATGATGATTTGATGATGCTTACAGCTTTGGGCGCGGCTAGG 690  
 Qy 3880 AGTGCGATGCTCTTATGGAAGTTAACTATTAAGTGCTTATGCTTACAGAGCA 3939

Db 689 ATTTCATGCTCTTATNNATTTAACTATTATNTGCTTTATNNGCCCTTANMANCA 630  
 Qy 3940 ATTTCGAGCGGAGCAGTGAAGAAAGTGAATATGATGAGAAACTCAACCTTAT 3999  
 Db 629 ATTTCCTGAGCGAGAGTGAATTAAGTGAATATGATGAGAAACTCAACCTTAT 570  
 Qy 4000 TTAGATCTGTAATGAAACCCAGTCTTGTATTAATTCCTGCTGTGAGAACTTTAA 4059  
 Db 569 TTAGATCTTAAATGAAACCCGTTTTTTTGTATTAATTCCTGCTGTGAGAACTTTAA 510  
 Qy 4060 GTTTCAAATACCCAGATGCTTCACTGCTCATATGCTGATTAATGAGCCAGAGACAG 4119  
 Db 509 GTTTCAAATACCCAGATGCTTCACTGCTCATATGCTGATTAATGAGCCAGAGACAG 450  
 Qy 4120 GCGGTGAGAGTTATCAACTGCTGAAGTGAAGTGGTCCCTGTGACCTTGAAGTATGC 4179  
 Db 449 NCGTGAAGAGTTATCAACTGCTGAAGTGAAGTGGTCCCTGTGACCTTGAAGTATGC 390  
 Qy 4180 CAGGCGTCCGCTCTGCTGAGATGAGTGAAGTGAAGTGGTCCCTGTGACCTTGAAGTATGC 4238  
 Db 389 CAGGCGTCCGCTCTGCTGAGATGAGTGAAGTGAAGTGGTCCCTGTGACCTTGAAGTATGC 330  
 Qy 4239 TTTTATTTCTGTTTCAAGCTTCTGATCTTATGAACTTTGCTGTGA 4287  
 Db 329 TTTTATTTCTGTTTCAAGCTTCTGATCTTATGAACTTTGCTGTGA 281  
 RESULT 10  
 BU624356/c 638 bp mRNA linear EST 23-SEP-2002  
 LOCUS BU624356 UI-H-FGI-bj3-1-22-0-UI.s1 NCI CGAP FGI Homo sapiens cDNA clone  
 DEFINITION UI-H-FGI-bj3-1-22-0-UI 3', mRNA sequence.  
 ACCESSION BU624356  
 VERSION BU624356.1 GI:23290571  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 REFERENCE 1 (bases 1 to 638)  
 AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.  
 TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
 Tumor Gene Index  
 JOURNAL Unpublished  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: cgapds-remail.nih.gov  
 Tissue Procurement: James Martin  
 cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa  
 cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa  
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa  
 Clone Distribution: Clone distribution information can be obtained  
 from Dr. M. Bento Soares, bento-soares@uiowa.edu  
 Seq primer: M13 FORWARD  
 POLYA=yes.  
 FEATURES  
 source location/Qualifiers  
 1..638  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="UI-H-FGI-bj3-1-22-0-UI"  
 /issue\_type="Cell lines"  
 /dev\_stage="adult"  
 /lab\_host="DH10B (Life Technologies)"  
 /clone\_lib="NCI-CGAP\_FGI"  
 /note="Organ: Enchondroma; Vector: pT73-Pac (Pharmacia)  
 with a modified polylinker; Site 1: EcoR I; Site 2: Not I;  
 NCI CGAP FGI is a normalized cDNA library obtained from a  
 pool of mRNA from 2 cell lines from Enchondroma tissues.  
 The library was constructed according to Bonaldo, Lennon  
 and Soares, Genome Research, 6:791-806, 1996. First strand  
 cDNA synthesis was primed with an oligo-dT primer  
 containing a Not I site. Double stranded cDNA was ligated



Db	421	GGGTCGCCCTGTGGAGGATGAGACTTCAGGCTCCATCATCACTTCAGTCAATTTT	480
Oy	4243	ATTTCTGTTCAAGCTTCTGACTTTATGACATTTGGCTGTGA	4287
Db	481	ATTTCTGTTCAAGCTTCTGACTTTATGAACTTTGGCTGTGA	525
RESULT 12			
LOCUS	B6E13975		
DEFINITION	B6E13975 RIKEN full-length enriched, 0 day neonate head Mus musculus cDNA clone 4832419101 5', mRNA sequence.		
ACCESSION	B6E13975		
VERSION	B6E13975.1		
KEYWORDS	EST.		
SOURCE	Mus musculus (house mouse)		
ORGANISM	Mus musculus		
REFERENCE	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 648)		
AUTHORS	Arakawa,T., Carninci,P., Fukuda,S., Furuno,M., Hanagaki,T., Hara,A., Hiramoto,K., Hori,F., Ishii,Y., Ito,M., Kawai,J., Komno,H., Kouda,M., Koya,S., Matsuyama,T., Miyazaki,A., Nomura,K., Ohno,M., Okazaki,Y., Okido,T., Saito,R., Sakai,C., Sakai,K., Sano,H., Sasakita,D., Shibata,K., Shingawa,A., Shirai,T., Sogabe,Y., Suzuki,H., Tagami,M., Tagawa,A., Takahashi,F., Takeda,Y., Tanaka,T., Toya,T., Muramatsu,M., and Hayashizaki,Y.		
TITLE	RIKEN Mouse ESTs (Arakawa,T., et al. 2001)		
JOURNAL	Unpublished		
COMMENT	Contact: Yoshihide Hayashizaki Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), Yokohama Institute The Institute of Physical and Chemical Research (RIKEN) 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan Tel: 81-45-503-9222 Fax: 81-45-503-9216 Email: genome-res@gscc.riken.go.jp/ URL: http://genome.gsc.riken.go.jp/ Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh,M., Komno,H., Okazaki,Y., Muramatsu,M., and Hayashizaki,Y. Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. Genome Res. 10 (10), 1617-1630 (2000) Wagui,K., Fujiwaka,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watahiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsunura,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kita,A., and Hayashizaki,Y. RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer. Genome Res. 10 (11), 1757-1771 (2000) Komno,H., Fukunishi,Y., Shibata,K., Itoh,M., Carninci,P., Sugahara,Y., and Hayashizaki,Y. Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001) Kondo,S., Shingawa,A., Saito,T., Kiyosawa,H., Yamanaaka,I., Aizawa,K., Fukuda,S., Hara,A., Itoh,M., Kawai,J., Shibata,K., and Hayashizaki,Y. Computational Analysis of Full-length Mouse cDNAs Compared with Human Genome Sequences. Mamm. Genome. 12, 673-677 (2001) Please visit our web site (http://genome.gsc.riken.go.jp) for further details. e mouse tissues.		
FEATURES	Location/Qualifiers		
SOURCE	1..648		
	/organism="Mus musculus"		
	/mol_type="mRNA"		
	/strain="C57BL/6J"		
	/db_xref="taxon:10090"		
	/clone="4832419101"		
	/sex="mixed"		
	/tissue_type="head"		
	/dev_stage="0 day neonate"		

```

/lab host="DHI0B"
/cjone_lib="RIKEN full-length enriched, 0 day neonate head"_lib=
/notes="Site_1: SalI; Site_2: BamH; cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN, Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. 1st strand cDNA was primed with a primer [5' GAGGAGAGAAGATCCACAGACTCTTTTTTTTTTTTTTVN 3'], cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. cDNA went through one round of normalization to Rot = 10.0 and subtraction to Rot = 100.0. Second strand cDNA was prepared with the primer adaptor of sequence [5' GAGGAGAGAATTTCGAGTTAAATAAATAATCCCCCCCCC 3']. cDNA was cloned into the XhoI and BamH sites. Vector: a modified Bluescript KS(+) after bulk excision from Lambda PUC I."
BASE COUNT      195 a    118 c    140 g    194 t     1 others
ORIGIN
Query Match      11.7%; Score 501.6; DB 10; Length 648;
Best Local Similarity 86.0%; Pred. No. 2e-100;
Matches 555; Conservative 0; Mismatches 90; Indels 0; Gaps 0;
QY 2585 AATCATCTTAATTAAGACTGA CTGACAATAGGCTTACAGAGTACCCTGAAAACCTTGAGTT 2644
Db 3 AATCTGTCTTAATTAAGACTGA CTGACAATAGGCAACGTAGAAAGTAAACAGAACAAATCTTAAGAGAT 62
QY 2645 TCTCATTTCCCTCCATTAATACAGTGCATCGGACGTGAAAGATCAATACATCGSCAATTTGGAG 2704
Db 63 TCTCTCTTCTCCCTCGAATACGGCTCAATGGCAGTAGAAAGTTCGAATACACACCAANTGGAG 122
QY 2705 ATGTTCTTGAGCTTCATCAATAGGCTTAGGCTCATTTGATTCGGATGCCATTATGGCTGTG 2764
Db 123 ATAATCTTGTTGCTCTCATCAATAGGCTTATCTTACATGATCCGAGATGCCATTATGATATGTG 182
QY 2765 GTGAACAGAA CATGATAAATTTTGCTCCAAATATTTACATTTTGGATTAATCTGACTAAAA 2822
Db 183 GTGAACAGAA CATGATAAATTTTGCTCCAAATATTTACATTTTGGATTAATCTGACTAAAC 242
QY 2825 AGAACAACATGACAGATTAATTTGAAAAGAAAAGCTCTTCAATTAATAGGGAAGSTTAC 2888
Db 243 AGAACAACATGACAGATTAATTTGAAAAGAAAAGGCCCTTTCATPATATGAGGCAAGSTTACC 302
QY 2885 AGAGAGAAGCTCTCTATACAGAGGAAGATGGCTCTTTCAGAGCTTTTGGCAATTATGACC 2944
Db 303 AAAGGAGAGCTTCTATACAGAGGAAGATGGCTCTTTCAGAGCTTTTGGGACATTTGACT 362
QY 2945 CTTCTGGAGAGCACTTGCTGTGCTGACTTTTGTGTTTAAAGTGTTCCTTGAAGCCGATCCTT 3004
Db 363 CTTCTGGAGAGCACTTGCTGTGCTGACTTTTGTGTTTAAAGTGTTCCTTGAAGCCGATTAAT 422
QY 3005 ACATAGATTAATGATCAGAAATGTGTTTACACAGAAATATACATTTGGCTTAAAGACATACGA 3066
Db 423 ATATATGATTAATGATCAGAAATGTGTTTACACAGAAATATATCTTGGCTTAAATGCAATATAGA 482
QY 3065 AATCCAACGGGATATTTTGGGATCCAGAAAGAGATTCATATAGAGCTTCAAGGTGGCA 3124
Db 483 AATTCATATGGGATATTTTGGGAGCCAGAAAGAGATTCACAGTGAACCTTCAAGGTGGCA 542
QY 3125 ATAAAGTCCAGTAAACCTTACAGCCTATATTTGATCTTCTCTCGGATATAGAAAGT 3184
Db 543 CCABAAGCCAGTAAACCTTACAGCCTATATTTGATCTTCTCTCGGATATAGAAAGT 602
QY 3185 ATCAGCCTTAACATTTGATGTGCAGAGTCTATCTCAATTTTGGAGT 3229
Db 603 ATCAGCCTTAATATGATGTACAAAGATCAATCAAGTTTGGAGAT 647

```



CB158337  
LOCUS CB158337 613 bp mRNA linear EST 29-JAN-2003  
DEFINITION K-EST0217654 L18POOL1n1 Homo sapiens cDNA clone L18POOL1n1-10-H06  
ACCESSION CB158337  
VERSION CB158337.1 GI:28143473  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 613)  
Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,  
Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and  
Kim,Y.S.  
21C Frontier Korean EST Project 2001  
JOURNAL Unpublished  
COMMENT Contact: Kim YS  
Genome Research Center  
Korea Research Institute of Bioscience & Biotechnology  
52 Boeun-dong Yuseong-gu, Daejeon 305-333, South Korea  
Tel: +82-42-860-4470  
Fax: +82-42-860-4409  
Email: yongsung@mail.kribb.re.kr  
Plate: 10 row: H column: 06  
High quality sequence stop: 613.  
Location/Qualifiers  
1..613  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="L18POOL1n1-10-H06"  
/cell\_line="SNU-354+Cho-CK+Cho1-CK+HLK-3"  
/lab\_host="Top10F"  
/clone\_lib="L18POOL1n1"  
/note="Organ: Liver; Vector: pTTT3-Pac; Site\_1: EcoRI;  
Site\_2: NotI; The library was contributed by the Soares  
Laboratory and it was constructed as described by Bonaldo,  
M.F., Lennon, G. and Soares, M.B. (1996), Genome Research  
6(9): 791-806. RNA was prepared from harvested cell  
culture."  
BASE COUNT 183 a 111 c 142 g 177 t  
ORIGIN  
Query Match 11.6%; Score 498; DB 14; Length 613;  
Best Local Similarity 99.8%; Pred. No. 1.2e-99;  
Matches 509; Conservative 0; Mismatches 0; Indels 1; Gaps 1;  
QY 3652 CTTCGTGCTGACAGCCAAATGCAATTAATTCGCAAAATGCTTTGATTTGCTATT 3711  
DB 104 CTTGCTGTGTACAGCCAAATGCAATTAATTCGCAAAATGCTTTGATTTGCTATT 163  
QY 3712 TGTGACCTCAATGTTGATTAATTAATGGAAGCTTCTGGGCTCTTGAAGACGAAGATCT 3771  
DB 164 TGTGACCTCAATGTTGATTAATTAATGGAAGCTTCTGGGCTCTTGAAGACGAAGATCT 223  
QY 3772 ATCCAAATCAAGAAAGCTTTGATTAATTAATTCGCTTAAAGAAATTAAGTATCTC 3831  
DB 224 ATCCAAATCAAGAAAGCTTTGATTAATTAATTCGCTTAAAGAAATTAAGTATCTC 283  
QY 3832 AATCATGTGATTTGATGATGTTGATCAAGCTTTCCGGCCCGGGTGAAGTGGCATGGCT 3891  
DB 284 AATCATGTGATTTGATGATGTTGATCAAGCTTTCCGGCCCGGGTGAAGTGGCATGGCT 343  
QY 3892 CTTATGGAAGTTAACTATTAATTAATTAATGATGCTTTGATGATGATGATGATGATG 3951  
DB 344 CTTATGGAAGTTAACTATTAATTAATTAATGATGCTTTGATGATGATGATGATGATG 403  
QY 3952 GAGACAGTGAAGAAATGGAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 4011  
DB 404 GAGACAGTGAAGAAATGGAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 463  
QY 4012 AATGAACCAAGTTTGTGTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 4071

DB 464 AATGAACCAAGTTTGTGTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 523  
QY 4072 CAAG-ATGCTCAGTGTCAATGATGATTAATTAATTAATTAATTAATTAATTAATTAAT 4130  
DB 524 CAAGATGCTCAGTGTCAATGATGATTAATTAATTAATTAATTAATTAATTAATTAAT 583  
QY 4131 TTACACTCTGGAAGTGAAGTGTCTCTCTG 4160  
DB 584 TTACACTCTGGAAGTGAAGTGTCTCTCTG 613  
RESULT 14  
BX472602 515 bp mRNA linear EST 12-JUN-2003  
LOCUS BX472602  
DEFINITION DKFZp686C02145\_r1 686 (synonym: hlcc3) Homo sapiens cDNA clone  
ACCESSION DKFZp686C02145.5, mRNA sequence.  
VERSION BX472602  
KEYWORDS BX472602.1 GI:31666864  
SOURCE EST.  
ORGANISM Homo sapiens (human)  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 515)  
Bloeker,H., Boecker,M., Mewes,H.W., Weill,B., Amid,C., Oeanger,A.,  
Fobbo,G., Han,M. and Wiemann,S.  
EST (Bloeker,H., Boecker,M., Mewes,H.W., Weill,B., Amid,C., et al.)  
JOURNAL Unpublished  
COMMENT Contact: Bloeker H  
MIPS  
Ingolstaedter Landstr.1, D-85764 Neuberg, Germany  
This is the 5' sequence of the clone insert  
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer  
Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;  
Braunschweig/Germany) within the cDNA sequencing consortium of the  
German Genome Project.  
a1 sequence also available.  
This clone (DKFZp686C02145) is available at the RZPD in Berlin.  
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059  
Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.  
Location/Qualifiers  
1..515  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="DKFZp686C02145"  
/tissue="human skeletal muscle"  
/dev\_stage="adult"  
/lab\_host="DH10B"  
/clone\_lib="686 (synonym: hlcc3)"  
/note="Vector: pTriplex2; Site\_1: SfiI; Site\_2: SfiI;  
cDNA-collection"  
BASE COUNT 130 a 142 c 120 g 123 t  
ORIGIN  
Query Match 11.6%; Score 496; DB 13; Length 515;  
Best Local Similarity 100.0%; Pred. No. 3.3e-99;  
Matches 496; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 AATGAGGGCCCAAGCGCTCTGACCGCGGCCCAAGCGCTCTGCTGTGACAGCGCGGCTG 60  
DB 20 AATGAGGGCCCAAGCGCTCTGACCGCGGCCCAAGCGCTCTGCTGTGACAGCGCGGCTG 79  
QY 61 GCCGTGCTCCCGGCGCTCGGTTTCTGTGACAGCGCGGCGGATCATGAGCCCGGAGGA 120  
DB 80 GCCGTGCTCCCGGCGCTCGGTTTCTGTGACAGCGCGGCGGATCATGAGCCCGGAGGA 139  
QY 121 AATGTGACTTATTTGGGCTGAGCTTTGGAACATGCTCTTCAACAGTATCTGTGAAGCG 180  
DB 140 AATGTGACTTATTTGGGCTGAGCTTTGGAACATGCTCTTCAACAGTATCTGTGAAGCG 199

QY	181	GAGCTGCTCAAGACAGCATCAAACTCACTGCTCTGCTCTGGAAGCAAGAGCTTT	240
Db	200	GAGCTGCTCAAGACAGCATCAAACTCACTGCTCTGCTCTGGAAGCAAGAGCTTT	259
QY	241	GAAGAAAGGCTCTTTTAAGACACTTACTCTTCCATCATCACTCTGTGAACAGTGCAGATTGAG	300
Db	260	GAAGAAAGGCTCTTTTAAGACACTTACTCTTCCATCATCACTCTGTGAACAGTGCAGATTGAG	319
QY	301	ATTATATGAGCTACGCTGTAAACCGAGAGTACCAGAGATGAGATTTTATTCCTATATAGTACC	360
Db	320	ATTATATGAGCTACGCTGTAAACCGAGAGTACCAGAGATGAGATTTTATTCCTATATAGTACC	379
QY	361	CGCTTATCATTTTGAGACCAAGAGAAATATCTGTCTTCATTCAAACAGACAAAGGCTTATAC	420
Db	380	CGCTTATCATTTTGAGACCAAGAGAAATATCTGTCTTCATTCAAACAGACAAAGGCTTATAC	439
QY	421	AAGCCAAAGCAAGAGTGAAGTTTGCATTTGTAACTCTTCTCGAGATTTTAAAGCCTTAC	480
Db	440	AAGCCAAAGCAAGAGTGAAGTTTGCATTTGTAACTCTTCTCGAGATTTTAAAGCCTTAC	499
QY	481	AAAACCTCTTTAACA	496
Db	500	AAAACCTCTTTAACA	515

RESULT 15					
BX370169					
	BX370169	919 bp	mRNA	linear	EST 08-MAY-2001
LOCUS	BX370169				
DEFINITION	Homo sapiens PLACENTA COR 25-NORMALIZED Homo sapiens cDNA clone CS01081YB18 5-PRIME, mRNA sequence.				

ACCESSION	BX370169	GI:30455806
VERSION	BX370169.1	
KEYWORDS	EST.	
SOURCE	Homo sapiens	
ORGANISM	Homo sapiens (human)	
	Homo sapiens.	

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT

Genoscope - Centre National de Séquençage  
BP 191 91006 EVRY cedex - France  
Email: [seqr@genoscope.cns.fr](mailto:seqr@genoscope.cns.fr) Web : [www.genoscope.cns.fr](http://www.genoscope.cns.fr)  
Library was constructed by Life Technologies, a division of  
Invitrogen. This sequence belongs to sequence cluster 3578.F For  
more information about this cluster, see  
<http://www.genoscope.cns.fr/>  
[cgf-bin/cluster.cgi?seq=CS0BAG038ZH04.CS03584.1&cluster=3578.F](http://cgf-bin/cluster.cgi?seq=CS0BAG038ZH04.CS03584.1&cluster=3578.F).  
Contact : Feng Liang Email : [liang@life tech.com](mailto:liang@life tech.com) URL :  
<http://fulllength.invitrogen.com/> Invitrogen Corporation 1600  
Faraday Avenue Genoscope sequence ID : CS0BAG038ZH04.CS03584.1.

FEATURES	SOURCE	LOCATION/Qualifiers
	1.	.919
		/organism="Homo sapiens"
		/mol_type="mRNA"
		/db_xref="taxon:9606"
		/clone="CSOD1081YB18"
		/tissue_type="PLACENTA COT 25-NORMALIZED"
		/clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
		/note="1st strand cDNA was primed with a NotI-cligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."
BASE COUNT	285 a	213 c 177 g 242 t
ORIGIN		2 others

Query Match	11.5%	Score 493.8	DB 13	Length 919
Best Local Similarity	94.4%	Pred. No. 1.1e-98		
Matches 577	Conservative 0	Mismatches 27	Indels 7	Gaps 6

722 TGAATTCTAAGACATTAAATGTTACCATTCACGGCAAGATATCATATGCGAAGCCAGTGA 781

Db	4	TGAATTTAAAGCCTTTAAATGGTATGCATCAACGGGAAGTATACATATGGGAAGCCAGTGA	63
Qy	782	AAGGAGCCGTAAAGCCTTACATTTTAACTTTATCCTTTGGGAAAAGAAAAATTTTA	841
Db	64	AAGGAGCCGTAAAGCCTTACATTTTAACTTTATCCTTTGGGAAAAGAAAAATTTTA	123
Qy	842	CAAAAACATTTAAGATMAATGGATCTGCAAACTCTCTTTTAAATGATGAAGAGATGAAAA	901
Db	124	CAAAAACATTTAAGATMAATGGATCTGCAAACTCTCTTTTAAATGATGAAGAGATGAAAA	183
Qy	902	ATGTAAATGAAATCTTCAAAATGACCTTTCGAATTAACCTGATCTATCTTCCCTGGACCA-	960
Db	184	ATGTAAATGAAATCTTCAAAATGACCTTTCGAATTAACCTGATCTATCTTCCCTGGACCA	243
Qy	961	GTAGAAATTTTAAACCAACAGTGCAGAAATCAGTTACAGATTTTCAAGAAATGTAAAGCACT	1020
Db	244	GTAGAAATTTTAAACCAACAGTGCAGAAATCAGTTACAGATTTTCAAGAAATGTAAAGCACT	303
Qy	1021	AAATGTGTTCTTCAAGCAACATGATTAATCATTTGAATTTTGGATTATACTA-CTGTCTT	1079
Db	304	AAATGTGTTCTTCAAGCAACATGATTAATCATTTGAATTTTGGATTATACTA-CTGTCTT	363
Qy	1080	GAAGCCATCTCTCAACTTCACAGCCA-CTGTGAAGTAACTCGTGCATGGCAACCAAC	1138
Db	364	GAAGCCATCTCTCAACTTCACAGCCA-CTGTGAAGTAACTCGTGCATGGCAACCAAC	423
Qy	1139	TGACTCTTGAAGAAAGAAATATATGTAGTCAATACAGTGACACAGAGAACTATATCTG	1198
Db	424	TGACTCTTGAAGAAAGAAATATATGTAGTCAATACAGTGACACAGAGAACTATATCTG	483
Qy	1199	AGTACTGAGACGGATCTPAACAGTGGAAATCAGAAAAATGAAAGCTGTTCAGAAAAATAAAT	1258
Db	484	AGTACTGAGACGGATCTPAACAGTGGAAATCAGAAAAATGAAAGCTGTTCAGAAAAATAAAT	543
Qy	1259	A-TACTGTCCCCCAAGTGAACCTT-TTAAGATTGAAT--CCCAATCTGAGAGATTCC	1314
Db	544	ATTACTGTCCCCCAAGTGAACCTT-TTAAGATTGAATTTCCCAACCTCGAGAGATTCC	603
Qy	1315	AGTAGAGCTACA 1325	
Db	604	AGGAGAGCTCA 614	

Search completed: January 16, 2004, 20:57:23  
Job time : 7735 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 16, 2004, 12:38:20 ; Search time 14986 Seconds

(without alignments)  
11702.894 Million cell updates/sec

Title: US-10-020-095-3

Perfect score: 4287

Sequence: 1 atgcagggccacgcgcctctcctttaggaacttgcctcga 4287

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 2888711 segs, 2045481386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-Processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl:\*  
1: gb\_ba:\*  
2: gb\_hcg:\*  
3: gb\_in:\*  
4: gb\_om:\*  
5: gb\_ov:\*  
6: gb\_pac:\*  
7: gb\_ph:\*  
8: gb\_pl:\*  
9: gb\_pr:\*  
10: gb\_ro:\*  
11: gb\_ses:\*  
12: gb\_sy:\*  
13: gb\_un:\*  
14: gb\_vl:\*  
15: em\_ba:\*  
16: em\_fun:\*  
17: em\_hum:\*  
18: em\_in:\*  
19: em\_mu:\*  
20: em\_om:\*  
21: em\_or:\*  
22: em\_ov:\*  
23: em\_pac:\*  
24: em\_ph:\*  
25: em\_pl:\*  
26: em\_ro:\*  
27: em\_ses:\*  
28: em\_un:\*  
29: em\_vl:\*  
30: em\_hcg\_hum:\*  
31: em\_hcg\_hiv:\*  
32: em\_hcg\_ocher:\*  
33: em\_hcg\_mus:\*  
34: em\_hcg\_pin:\*  
35: em\_hcg\_rnd:\*  
36: em\_hcg\_tam:\*  
37: em\_hcg\_vrt:\*  
38: em\_sy:\*  
39: em\_higo\_hum:\*  
40: em\_higo\_mus:\*  
41: em\_higo\_other:\*

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	4224.4	98.5	4761	6	AX534934
2	4224.4	98.5	5883	6	AF410459 Homo sapi
3	4224.4	98.5	5895	6	AX534938
4	4222.8	98.5	4761	6	AX534936
5	4222.8	98.5	5895	6	AX534940
6	4221.4	98.5	4335	6	AX537612
7	4219.8	98.4	4335	6	AX537612
8	3187.4	74.4	3535	6	AX534942
9	3185.8	74.3	3535	6	AX534944
10	2709.2	63.2	4735	10	AY083458
11	2709.2	63.2	5644	10	BC052443
12	2448.2	57.1	2938	9	HS0805581
13	1969.4	45.9	2273	9	AK095888
14	730.8	17.0	1300	6	AX083146
15	232	5.4	163577	9	AL590428
16	230	5.4	176144	2	AC012408
17	219.6	5.1	5603	3	CIN431688
18	203.8	4.8	4678	3	D83196
19	182	4.2	251250	2	AC097023
20	177.4	4.1	91419	9	AL591480
21	175	4.1	149597	2	AC034271
22	174	4.1	4527	6	AX054953
23	174	4.1	4560	6	AX054954
24	171.6	4.0	2629	3	AY118302
25	167.6	3.9	4715	5	GG0VOM
26	167	3.9	4578	3	AY122084
27	167	3.9	4579	3	DME269539
28	160	3.7	144589	2	AC112668
29	160	3.7	191258	2	AC118254
30	155.4	3.6	4620	10	RNA113
31	154.8	3.6	176144	2	AC012408
32	154.2	3.6	4636	10	RAT113A
33	152.2	3.6	4220	3	DME269538
34	151.2	3.5	4575	10	GPIMSPB
35	149.8	3.5	2501	10	RAT113
36	146.4	3.4	4950	5	XELENDO
37	143.8	3.4	122013	3	AC092395
38	143.8	3.4	164414	3	AC092395
39	143.8	3.4	168503	3	AC115484
40	143.8	3.4	275699	3	AE003649
41	143.8	3.4	337178	3	DROSADH08
42	140.8	3.3	191258	2	AC118254
43	136	3.2	4623	10	BC036299
44	136	3.2	4672	10	BC051037
45	134.8	3.1	2689	5	AB026130

# ALIGNMENTS

RESULT 1  
AX534934  
LOCUS AX534934 4761 bp DNA linear PAT 22-NOV-2002  
DEFINITION Sequence 1 from Patent WO02070696.  
ACCESSION AX534934  
VERSION AX534934.1 GI:25261474  
KEYWORDS  
SOURCE  
ORGANISM Homo sapiens (human)  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE  
1 Schuh, A. and Sutherland, R. D.  
Cd109 nucleic acid molecules polypeptides and methods of use  
JOURNAL Patent: WO 02070696-A 1 12-SEP-2002;

Pred. No. is the number of results predicted by chance to have a

FEATURES  
source Schuh, Andre (CA) ; Sutherland, Robert D. (CA)

## CDS

1..4761  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
113..4450  
/note="unnamed protein product"  
/codon\_start=1  
/protein\_id="CAD57231.1"  
/db\_xref="gi:25261475"  
/translation="MQGPPLITAAHLICVCTAALAAVAPRPFLVTAAGIIRPGSNVTL  
GVELLEHCPQVTVKAEILLKTAASNLTVSLAEAEVFEKSPKTLITLSPILNSADELY  
ELRYTRGRDEILFNSNTRLSFEKRIISVFIOTKALYKPKQEKFPFIVLFSDFKRY  
KTSNIIILKDEKSNLIQOMLSQSDGLVTSKTFOLDSPHILGDMISIQVNDQYYS  
FOSEYVLPKFEVTLQPLCYCSNKKHNGITAKYTKYGVKVDLITLTPFSFWK  
KKNITKPKINGSANFSPNDEMKVNSNGSLSEYLDSPRGVELITLTPYSVTRI  
SRNASTNVFKQHDITLIEFDYITVLKSLNFTAVKTRADGQQLLEERKNVIT  
VTORNYTEWSGNSGNQKMEAVOKINVTPOSGTFKIEPILDSSELQKAVFLGS  
KSMASVHSLFKSPKTYIQLKRDENIKVSGPFLVSGNKRLELSYMVNSQOLVA  
VKONSTWFSLLTPENSTPKACIVVYIEDGEIISDPLKIPOLVFNKIKLWMSYV  
KAPSEKSLRISVTOPDSIVGIYAVDKSVLMAASNDITMENVTHELKLYNGSYLG  
MPNMSRAVQECGLMLTDLNATDYIDGVINDAEYARFMEENEGHIVIDHDSLS  
SPHVRKHPETWLDWNGRYRIQEFVTVPDSITSWATGFIISDILGGLTTTV  
ELQAFQPFILNLPYSVIRGEBALETTIFNYLKATEVYIIEKSDKDIIMTSE  
INATGHOQTLVPSDEGATVLPFRPHTLGEIPITVALSPASDAVTOMILVABEII  
EKXSOSILDLNRLQSTLKTLSFSPNVTGSEBOITLADVLGSPINGLASLI  
IRMPYGGGEONINPAPNIIYLDYLTKKKOLLTKKALSPMGQYORLELYORENG  
SPSAFGNVDPSGTSWLSFVLRCLFLEADPIDIDQNLHRTYTLKHGKSNENP  
GRVHSELQGNKSPVLTITAVITSLGKRYQPNIDVQESITHELESEFGSISDNT  
LALITVALSVSGSPKAKELMLTWRAEGEGMQFVWSSEKSDSQPRLDIEVA  
YALISHLQFQTSBGIPIMRWLSRQNSLQGFASDTOTVVALKALSFALMETERN  
IOVTYVSPSSPVKFLIDHNRLLQTELAVALVQPAVNIANSANGFALCOLAVNV  
IKAGSSRRRSIONOEAFDLDAVKNKDILNVDINVCISFSGPGRSGMALMENVL  
LSGPMVSEASISLSEYTKVYEDGKANLYIDSYNEFPCNITPAVNPKNYSYQDS  
VSIYDYEPRQAVRSINSEVKSCLDQVQCRPCEDSAGSHHSVIFLFCFK  
LILYFMEIWL"

BASE COUNT 1420 a 912 c 1016 g 1413 t  
ORIGIN

Query Match 98.5%; Score 4224.4; DB 6; Length 4761;  
Best Local Similarity 98.8%; Pred. No. 0; 4450  
Matches 4286; Conservative 0; Mismatches 1; Indels 51; Gaps 1;

1 ATGCAAGGCGCCACGCTCTGACCGCGCGCCACGCTCTGCGTGTGACCGCGCGCTG 60  
113 ATGCAAGGCGCCACGCTCTGACCGCGCGCCACGCTCTGCGTGTGACCGCGCGCTG 172  
61 GCCGTGCTCCCGGCTCGGTTTCTGGTGAACAGCCCGAGGATCATCAGGCCCGAGA 120  
173 GCCGTGCTCCCGGCTCGGTTTCTGGTGAACAGCCCGAGGATCATCAGGCCCGAGA 232  
121 AATGTGACTATTTGGGTGAGCTTCTGGAACACTGCGCTTCAACAGGTGACTGTGAAGCG 180  
233 AATGTGACTATTTGGGTGAGCTTCTGGAACACTGCGCTTCAACAGGTGACTGTGAAGCG 292  
181 GAGTGTCTCAAGACAGCATCAACCTGCTGCTGCTCGAAGAGAGAGAGTCTTT 240  
293 GAGTGTCTCAAGACAGCATCAACCTGCTGCTGCTCGAAGAGAGAGAGTCTTT 352  
241 GAAAAAGGCTCTTTTAAGACACTTACTTCCATCACTACTGAAACAGTGCAGATGAG 300  
353 GAAAAAGGCTCTTTTAAGACACTTACTTCCATCACTACTGAAACAGTGCAGATGAG 412  
301 ATTATGAGCTAGCTGTAAACCGACGTACCCAGAGATGAGATTTATTTCTTAATAGTACC 360  
413 ATTATGAGCTAGCTGTAAACCGACGTACCCAGAGATGAGATTTATTTCTTAATAGTACC 472  
361 CGCTTATCATTTGAGACCAAGAGAAATCTGCTTCAATTCAACAGACAGGCTTTATAC 420  
473 CGCTTATCATTTGAGACCAAGAGAAATCTGCTTCAATTCAACAGACAGGCTTTATAC 532  
421 AAGCAAGCAAGAGAGTTCGCAATGTTACACTCTTCTCAGATTTTAAAGCTTTAC 480

533 AAGCAAGCAAGAGAGTTCGCAATGTTACACTCTTCTCAGATTTTAAAGCTTTAC 592  
481 AAAACCTCTTTAAACAATTCCTCACTTAAGAGACCCCAAAATCAAATTTGATCCAAAGTGGTTG 540  
593 AAAACCTCTTTAAACAATTCCTCACTTAAGAGACCCCAAAATCAAATTTGATCCAAAGTGGTTG 652  
541 TCACAACAAAGATATCTTGAGACATTTCCAAACCTTTCAAGCTATCTTCCATCCAAATA 600  
653 TCACAACAAAGATATCTTGAGACATTTCCAAACCTTTCAAGCTATCTTCCATCCAAATA 712  
601 CTTGCTGACCTGCTCTATTTCAAGTTCAAGTGAATGACCAACATATTAATCAATTTTACAG 660  
713 CTTGCTGACCTGCTCTATTTCAAGTTCAAGTGAATGACCAACATATTAATCAATTTTACAG 722  
661 GTTTCGAATATGATATTAACAAATTTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 720  
773 GTTTCGAATATGATATTAACAAATTTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 832  
721 ATGAATTTCTAAGCATTTTAAATGATGATACATCAACGCAAGATATCATATGGAAGCAGATG 780  
833 ATGAATTTCTAAGCATTTTAAATGATGATACATCAACGCAAGATATCATATGGAAGCAGATG 892  
781 AAAGAGACGTAAACGCTTACATTTTACCTTTATCTTTTGGGAAAGAAATAAT 840  
893 AAAGAGACGTAAACGCTTACATTTTACCTTTATCTTTTGGGAAAGAAATAAT 952  
841 ACAAACATTTTAAATGATATGATATGATATGATATGATATGATATGATATGATATGATATGATATG 900  
953 ACAAACATTTTAAATGATATGATATGATATGATATGATATGATATGATATGATATGATATGATATG 1012  
901 AATGTAATGATATCTTCAATGATGATCTTCTGATATCTTCTGATATCTTCTGATATCTTCTGATAT 960  
1013 AATGTAATGATATCTTCAATGATGATCTTCTGATATCTTCTGATATCTTCTGATATCTTCTGATAT 1072  
961 GTGAAATTTTAAACCAAGTGAACGAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1020  
1073 GTGAAATTTTAAACCAAGTGAACGAATGATGATGATGATGATGATGATGATGATGATGATGAT 1132  
1021 AATGTGCTCTTCAAGCAACATGATATCATCATGATGATGATGATGATGATGATGATGATGATGATG 1080  
1133 AATGTGCTCTTCAAGCAACATGATATCATCATGATGATGATGATGATGATGATGATGATGATGATG 1192  
1081 AAGCCATCTCTCAACTTCAAGCAACATGATATCATCATGATGATGATGATGATGATGATGATGATG 1140  
1193 AAGCCATCTCTCAACTTCAAGCAACATGATATCATCATGATGATGATGATGATGATGATGATGATG 1252  
1141 ACTCTTGAGAAAGAAATTAATGATGATATGATGATATGATGATATGATGATATGATGATATGATG 1200  
1253 ACTCTTGAGAAAGAAATTAATGATGATATGATGATATGATGATATGATGATATGATGATATGATG 1312  
1201 TACTGAGAGGGAATCTTAAAGTGAATCAAGAAATGGAAGCTGTGAGAAATTAATTAAT 1260  
1313 TACTGAGAGGGAATCTTAAAGTGAATCAAGAAATGGAAGCTGTGAGAAATTAATTAAT 1372  
1261 ACTGTCCCCCAAGTGAATCTTTAAGATTTGAATTTCCCAATCTCTGAGGATTTCCATGAG 1320  
1373 ACTGTCCCCCAAGTGAATCTTTAAGATTTGAATTTCCCAATCTCTGAGGATTTCCATGAG 1432  
1321 CTACAGTTGAAGGCTATTTCTTGGTATGAAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1380  
1433 CTACAGTTGAAGGCTATTTCTTGGTATGAAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1492  
1381 AAGTCTCTTGAAGTGAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAAT 1440  
1493 AAGTCTCTTGAAGTGAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAAT 1552  
1441 TCAGCTTTTGAAGTGTGTTAGTGAAGCAAAAGATTTGAAGGATTTAAGCTATATGAT 1500  
1553 TCAGCTTTTGAAGTGTGTTAGTGAAGCAAAAGATTTGAAGGATTTAAGCTATATGAT 1612  
1501 GTATCCAGGGAACAGTTGAGCTGTAGAGAAACAAATTTCAACAAATGTTCTTTTAA 1560

Db 1613 GATCAGGGGACAGTTGGCTGTAGAAAACAAATTCACAATGCTTCTTTACA 1672  
 Oy 1561 CCAGAAAATTCCTGGACCTCCAAAAGCCTGTGTAAATTGTATTAATTAAGATGTGG 1620  
 Db 1673 CCAGAAAATTCCTGGACCTCCAAAAGCCTGTGTAAATTGTATTAATTAAGATGTGG 1732  
 Oy 1621 GAAATTAATAGTGTCTTAAATAATCCTGTTCACTGTTGTTTAAATAAGTAAG 1680  
 Db 1733 GAAATTAATAGTGTCTTAAATAATCCTGTTCACTGTTGTTTAAATAAGTAAG 1792  
 Oy 1681 CTAATTTGGAAGTAAGTGAACCTGACCACTGGAAGGCTCTCTTGAATCTCTGG 1740  
 Db 1793 CTAATTTGGAAGTAAGTGAACCTGACCACTGGAAGGCTCTCTTGAATCTCTGG 1852  
 Oy 1741 ACACAGCCTGATCTCATAGTGGATTTAGCTGTGTAACAAAAGTGTGAATCTGAAT 1800  
 Db 1853 ACACAGCCTGATCTCATAGTGGATTTAGCTGTGTAACAAAAGTGTGAATCTGAAT 1912  
 Oy 1801 GCCTTAATGATTAATCAATGGAATAATGTGTCCATGAGTTGGAATCTTATAACACAGA 1860  
 Db 1913 GCCTTAATGATTAATCAATGGAATAATGTGTCCATGAGTTGGAATCTTATAACACAGA 1972  
 Oy 1861 TATTATTTAGGATGTTCATGAATCTTTTGCAGTCTTTACAGAAATGTGACTGTGGTA 1920  
 Db 1973 TATTATTTAGGATGTTCATGAATCTTTTGCAGTCTTTACAGAAATGTGACTGTGGTA 2032  
 Oy 1921 TTGACAGATGCAAACTCCAGAAAGATTATTTGATGTTTATGACAAATGACAAATAT 1980  
 Db 2033 TTGACAGATGCAAACTCCAGAAAGATTATTTGATGTTTATGACAAATGACAAATAT 2092  
 Oy 1981 GCTGAGAGTTTATGAGGAAAATGAGAGACATAATTTGATATTCATGACTTTCTTGG 2040  
 Db 2093 GCTGAGAGTTTATGAGGAAAATGAGAGACATAATTTGATATTCATGACTTTCTTGG 2152  
 Oy 2041 GGTACAGTCCACATGTCCGAAAGATTTTCCAGAGCTTGGATTTGGCTAGACACACAC 2100  
 Db 2153 GGTACAGTCCACATGTCCGAAAGATTTTCCAGAGCTTGGATTTGGCTAGACACACAC 2212  
 Oy 2101 ATGGGTTACAGATTTTACCAAGATTTGAAATCTGTACCTGATCTTCTACTCTTGG 2160  
 Db 2213 ATGGGTTACAGATTTTACCAAGATTTGAAATCTGTACCTGATCTTCTACTCTTGG 2272  
 Oy 2161 GTGGCTACTGTTTGTGATCTCTGAGACCTGGGCTTTGACCTAACACTACTCCAGTG 2220  
 Db 2273 GTGGCTACTGTTTGTGATCTCTGAGACCTGGGCTTTGACCTAACACTACTCCAGTG 2332  
 Oy 2221 GAGCTCAAGCCTTCCAACTTTTTCATTTTGTGAATCTTCCCTACTCTGTATCAGA 2280  
 Db 2333 GAGCTCAAGCCTTCCAACTTTTTCATTTTGTGAATCTTCCCTACTCTGTATCAGA 2392  
 Oy 2281 GGTGAAGATTTGCTTTGAAAATACTAATTAATTAATTTGAAAAGATGCCACTGAGTT 2340  
 Db 2393 GGTGAAGATTTGCTTTGAAAATACTAATTAATTTGAAAAGATGCCACTGAGTT 2452  
 Oy 2341 AAGTAATCAATGAGAAAAGTGAACAATTTGATTAATCTAATGACTTCAAGTGAATAAT 2400  
 Db 2453 AAGTAATCAATGAGAAAAGTGAACAATTTGATTAATCTAATGACTTCAAGTGAATAAT 2512  
 Oy 2401 GCCACAGGCGACCAAGACACCTTGTGTTCCAGTGAAGATGGGCAACTGTCTTTT 2460  
 Db 2513 GCCACAGGCGACCAAGACACCTTGTGTTCCAGTGAAGATGGGCAACTGTCTTTT 2572  
 Oy 2461 CCCATCAGGCGCAACATCTGGAGAAAATTTCTATTCACAGTCCACAGCTTTTCAACCACT 2520  
 Db 2573 CCCATCAGGCGCAACATCTGGAGAAAATTTCTATTCACAGTCCACAGCTTTTCAACCACT 2632  
 Oy 2521 GCTTCTGATGCTGTCAACCAAGATGATTTAGTAAAGCTGAAGAAATGAATAATCAAT 2580  
 Db 2633 GCTTCTGATGCTGTCAACCAAGATGATTTAGTAAAGCTGAAGAAATGAATAATCAAT 2692  
 Oy 2581 TCACAATCATCTTATTTAGACTTGAACAAATAGCTACAGATACCTTGAAAACTTTG 2640  
 Db 2693 TCACAATCATCTTATTTAGACTTGAACAAATAGCTACAGATACCTTGAAAACTTTG 2752

Oy 2641 AGTTTCTCATTTCTCTCTAATAACAGTGAAGTGGCAGTGAAGAGTTCAGATCACTGCAATT 2700  
 Db 2753 AGTTTCTCATTTCTCTCTAATAACAGTGAAGTGGCAGTGAAGAGTTCAGATCACTGCAATT 2812  
 Oy 2701 GGAGATGTTCTTGGTCTTCCATCAATGAGCTTACCTCATTTGATTTGGATGCTTATGGC 2760  
 Db 2813 GGAGATGTTCTTGGTCTTCCATCAATGAGCTTACCTCATTTGATTTGGATGCTTATGGC 2872  
 Oy 2761 TGTGTGAACAGAACATATAATTTTGTCTCCAAATATTTCAATTTGATTAATCTGACT 2820  
 Db 2873 TGTGTGAACAGAACATATAATTTTGTCTCCAAATATTTCAATTTGATTAATCTGACT 2932  
 Oy 2821 AAAAAAACAACCTGACAGATTAATTTGAAGAAAAAGCTTTTCAATTTATGAGCAAGT 2880  
 Db 2933 AAAAAAACAACCTGACAGATTAATTTGAAGAAAAAGCTTTTCAATTTATGAGCAAGT 2992  
 Oy 2881 TACAGAGAGAACTTCTCTATCAGAGGGAATGCTCTTCACTGCTTTTGGAAATAT 2940  
 Db 2993 TACAGAGAGAACTTCTCTATCAGAGGGAATGCTCTTCACTGCTTTTGGAAATAT 3052  
 Oy 2941 GACCTTTCTGGAGACCTTGGTGTCTGACGCTTTGTTTAAAGTGTCTTGAAGCGAT 3000  
 Db 3053 GACCTTTCTGGAGACCTTGGTGTCTGACGCTTTGTTTAAAGTGTCTTGAAGCGAT 3112  
 Oy 3001 CTTTACATAGATTAATGATCAGAAATGTGTACACAGAAATCACTGCTTAAAGAGAT 3060  
 Db 3113 CTTTACATAGATTAATGATCAGAAATGTGTACACAGAAATCACTGCTTAAAGAGAT 3172  
 Oy 3061 CAGAAATCCAAACGATGAATTTTGGATCCAGAAAGATGATTCATAGTGAAGCTTCAAGT 3120  
 Db 3173 CAGAAATCCAAACGATGAATTTTGGATCCAGAAAGATGATTCATAGTGAAGCTTCAAGT 3232  
 Oy 3121 GGCAATAAAGTCCAGTAACTTACAGCTTATATTTGATCTTCTCTGGGAATATGA 3180  
 Db 3233 GGCAATAAAGTCCAGTAACTTACAGCTTATATTTGATCTTCTCTGGGAATATGA 3292  
 Oy 3181 AAGTATCAAGCTTAATGATGATGTCAGAGCTTATCCATTTTGTGAGTCTGAATTCAGT 3240  
 Db 3293 AAGTATCAAGCTTAATGATGATGTCAGAGCTTATCCATTTTGTGAGTCTGAATTCAGT 3352  
 Oy 3241 AGAGGAATTTCAACAATTAATCTCTAGCCCTTAACTTAATGCAATTTGATCACTGAGGG 3300  
 Db 3353 AGAGGAATTTCAACAATTAATCTCTAGCCCTTAACTTAATGCAATTTGATCACTGAGGG 3412  
 Oy 3301 AGTCTTAAGGGAAGAGCTTTGAAATGTCTGACTTGAAGACAGAAACAGAGTGGC 3360  
 Db 3413 AGTCTTAAGGGAAGAGCTTTGAAATGTCTGACTTGAAGAGCAGAAACAGAGTGGC 3472  
 Oy 3361 ATGCAATCTGGGTGTATCAGAGTCCAAACTTTCTGACTCTGGACAGCAAGCTCCCTG 3420  
 Db 3473 ATGCAATCTGGGTGTATCAGAGTCCAAACTTTCTGACTCTGGACAGCAAGCTCCCTG 3532  
 Oy 3421 GATATTTAAGTTTGCAGCTTATGACTGTCTCACAACCTTCAATTTCAAGCTTCTGAG 3480  
 Db 3533 GATATTTAAGTTTGCAGCTTATGACTGTCTCACAACCTTCAATTTCAAGCTTCTGAG 3592  
 Oy 3481 GGAATCCCAATTTAGAGTGTGCTAAGCAGGCAAGAAATTAAGCTTGGTGTTCATCT 3540  
 Db 3593 GGAATCCCAATTTAGAGTGTGCTAAGCAGGCAAGAAATTAAGCTTGGTGTTCATCT 3652  
 Oy 3541 ACTCAGATTAACACTGTGGCTTTAAAGGCTGTGTCTGAATTTGAGCCCTTAATGAATCA 3600  
 Db 3653 ACTCAGATTAACACTGTGGCTTTAAAGGCTGTGTCTGAATTTGAGCCCTTAATGAATCA 3712  
 Oy 3601 GAAAGGCAAAATATCAAGTACCGTGAAGGGGCTTACTCAACAAATGCTTCTTGAAGTTT 3651  
 Db 3713 GAAAGGCAAAATATCAAGTACCGTGAAGGGGCTTACTCAACAAATGCTTCTTGAAGTTT 3772  
 Oy 3652 -----CTTGTGTGATCAGGCA 3669  
 Db 3773 CTGATTGACACACAAACCGTTACTCTTGCAGACAGAGCTTGTGTGTACAGGCA 3832

```

QY 3670 ATGGCAGTAAATATTTTCGCAAAATGGTTTGGATTTGCTATTTGTCAGCTCAATGTTGTA 3729
DB 3833 ATGCAGTAAATATTTTCGCAAAATGGTTTGGATTTGCTATTTGTCAGCTCAATGTTGTA 3892
QY 3730 TATATATGTAAGGCTTTGGGCTTTCTAGAAAGAGATTTATCCAAATCAAGAAAGCC 3789
DB 3893 TATATATGTAAGGCTTTGGGCTTTCTAGAAAGAGATTTATCCAAATCAAGAAAGCC 3952
QY 3790 TTTGATTTAGATGTTGCTGTAAGAAAGAAATTAAGATGATCTCAATCATGTTGAT 3849
DB 3953 TTTGATTTAGATGTTGCTGTAAGAAAGAAATTAAGATGATCTCAATCATGTTGAT 4012
QY 3850 GTGTGTACAGCTTTTGGGCGCGGATGAGTGGCATGCTCTTATGAAAGTTAACTTA 3909
DB 4013 GTGTGTACAGCTTTTGGGCGCGGATGAGTGGCATGCTCTTATGAAAGTTAACTTA 4072
QY 3910 TTAAGTGGCTTTATGAGTGGCTTTCAAGAGCAATTTCTGAGCGGAGACAGTAAAGAGTG 3969
DB 4073 TTAAGTGGCTTTATGAGTGGCTTTCAAGAGCAATTTCTGAGCGGAGACAGTAAAGAGTG 4132
QY 3970 GAATATGATCATGGAAGAACTCAACCTCTATTAGATTTCTGTAAGTAAAGAAAGCTTTGT 4029
DB 4133 GAATATGATCATGGAAGAACTCAACCTCTATTAGATTTCTGTAAGTAAAGAAAGCTTTGT 4192
QY 4030 GTTATATTTCTGCTGTGAGAACTTTAAAGTTCAATACCAGATGCTTCAGTGTCC 4089
DB 4193 GTTATATTTCTGCTGTGAGAACTTTAAAGTTCAATACCAGATGCTTCAGTGTCC 4252
QY 4090 ATAGTGGATTACTATGAGCCCAAGAGACAGCGCGGTGAGAAATCAACTCTGAAGTGAAG 4149
DB 4253 ATAGTGGATTACTATGAGCCCAAGAGACAGCGCGGTGAGAAATCAACTCTGAAGTGAAG 4312
QY 4150 CTGTCTCTCTGAGACCTTTGAGTATGATGATGATGATGATGATGATGATGATGATGATGAT 4209
DB 4313 CTGTCTCTCTGAGACCTTTGAGTATGATGATGATGATGATGATGATGATGATGATGATGAT 4372
QY 4210 TCAGAGCTCCCATCATCACTCTTCAGTCACTTTTATTTTCTGTTCAAGCTTCTGACTTT 4269
DB 4373 TCAGAGCTCCCATCATCACTCTTCAGTCACTTTTATTTTCTGTTCAAGCTTCTGACTTT 4432
QY 4270 ATGGAACCTTTGGCTGTGA 4287
DB 4433 ATGGAACCTTTGGCTGTGA 4450

```

**RESULT 2**  
AF410459

**LOCUS** AF410459 5883 bp mRNA linear PRI 02-MAR-2002

**DEFINITION** Homo sapiens CD109 (CD109) mRNA, complete cds.

**ACCESSION** AF410459

**VERSION** AF410459.1 GI:19071208

**KEYWORDS**

**SOURCE** Homo sapiens (human)

**ORGANISM** Homo sapiens

**REFERENCE** Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

**AUTHORS** Lin, M., Sutherland, D.R., Horsfall, W., Toty, N., Yeo, E., Nayar, R., Wu, X.F., and Schuh, A.C.

**TITLE** Cell surface antigen CD109 is a novel member of the alpha(2) macroglobulin/C3, C4, C5 family of thioester-containing proteins

**JOURNAL** Blood 99 (5), 1683-1691 (2002)

**PubMed** 11861284

**REFERENCE** 2 (bases 1 to 5883)

**AUTHORS** Lin, M., Sutherland, D.R., Horsfall, W., Toty, N., Yeo, E., Nayar, R., Wu, X.F., and Schuh, A.C.

**TITLE** Direct Submission

**JOURNAL** Submitted (14-AUG-2001) Medicine, University of Toronto, 1 King's College Circle, Room 7366, Toronto, Ontario M5S 1A8, Canada

**FEATURES** 1..5883

**source** /organism="Homo sapiens"

```

/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="K1"
1..5883
/gene="CD109"
1..112
/gene="CD109"
113..4450
/gene="CD109"
/note="associated with the Gov alloantigen system"
/codon_start=1
/protein_id="AA184159.1"
/db_xref="GI:19071209"
/translation="MOGPPLTAAHLICVCTALAAVAPGPPLVTAAGLIRPGAVTI
GVELLHCSQVTVKAELEKTAENLTVSVLEAGVEKSGFKTLTLPSPLENDREY
ELAVTRTODELIFNSSTLSPTKRLSVIQLDPAKYKPKVKRYVTLPEDFKY
KISLNLTKDPSNLIQOMLSQSDGLVSKITQLSHPILGMSIOVNDOTYOS
FOVSEYVLPEFVTLQTPLYCSMNSRLHNGITAKTYGKPVAGVTLPLPSFWGK
KNITTKTFKINGSANFSFNDKEMKNVDSNGLEVLDSPPVEILTVTESVGI
SRVSTNVFPEKODYILPEFDTTVLKPSLNPAVYKRLPADGNOLTEERRNVVT
VQORNTYVMSGSGNOMKAEVOKINVTYVPSGTFKIEPILDESSLOLKAFLPS
KSMAYHSLFKPSKTYTQKTRDENIKOSPELVVSGKRLKELSYRVSGQLVA
VGKQNSTMSFLPENSWTPKACVIVYIIDDGELISDVLKIPQVLFPKNIKLYMSKY
KAPSEKSVLSRLSVTPDPSIVGIYAVDKSVNLMAANDITMENVAHELINTGYLG
MFNNSFAVPECEGLAVLTDLNLTQDYIDGVYDAEYAEERMEENEGHIVIDFSLGS
SPHVRKHPEPTWIMDNGYRIYQSEFVETVPSISWATGVISDGLGLTTPV
ELQAPQPFITFLNPSYVIRGEFALEITFVNTLKQATRYKVIIEKSDKEDILMNE
INATGHOQLTVLSEDAVTLFIRPTHELEIYITTTALSPTRASDVTOMIKACBI
EKYSQSLDLTDLNRQLTKTSLSPENVTGSEVQVTAIGDVLGSLNGLSL
IRMPYCGEONMNFAPNIIYLDYLRKQKLTDLNEKALSPKQORSLYQREBQ
SFAPGNYPDSGSTWLSAFVLRCEFLADPEYIDIDQVNLHRTYMLKGHOSNCEPMP
GRVHSELQGNKSPVTLTAIVTSLGKRYKQPNIDVQSHIFLSEFRRGISDNYT
LALITVALSVGSPKAKELNMLTWRABECCGMPVSSKXLSDSMOPSLIEVVA
YALSHLPQRTSEBGPIMKLSRQNSLGRASDQDTTALALSEFALNMTERN
IQVTVSPSSPSPVKFLIDHNRLLIQTELELAVQPAVVISANGFEALCOLAVNL
VKAGSSRRRSRIONQEPDLDAVENKEDLHNVLDLVNLSGSGRSMALMEVNL
LSGFVPEBAILSLSEYKVEYDHGKNTLIDVNESTQFVNI.PAVANFVSTODAS
VSLVDYVPEPRQAVRSYNSVVKLSQDLGSDVQGRCEBAGSGSHHSVIFCFK
LTFMELMT."
4451..5883
/gene="CD109"
4748
/misc_feature
/gene="CD109"
/note="alternative polyA site found on clone K1"
1697 a 1173 c 1241 g 1772 t
BASE COUNT
ORIGIN
Query Match 98.5%; Score 4224.4; DB 9; Length 5883;
Best Local Similarity 98.8%; Pred. No. 0;
Matches 4286; Conservative 0; Mismatches 1; Indels 51; Gaps 1;

```

QY 1 ATGAGGAGCCACCGCTCTGACCGCGCCACCTCTCTGCGTGTGACCGCGCGCTG 60

DB 113 ATGAGGAGCCACCGCTCTGACCGCGCCACCTCTCTGCGTGTGACCGCGCGCTG 172

QY 61 GCGTGGCTCCGGGCGCTGCTTGTGTGAGACGCGCCAGAGATCATCAGCCCGAGGA 120

DB 173 GCGTGGCTCCGGGCGCTGCTTGTGTGAGACGCGCCAGAGATCATCAGCCCGAGGA 222

QY 121 AATGTGACTATTTGGGAGTGTGAGAACTGACCTTCAAGAGTGAAGAGCG 180

DB 233 AATGTGACTATTTGGGAGTGTGAGAACTGACCTTCAAGAGTGAAGAGCG 292

QY 181 GAGCTGCTCAAGACAGATCAAACTGCTGTCTGTCTGTGAAGCAGAGAGTCTTT 240

DB 293 GAGCTGCTCAAGACAGATCAAACTGCTGTCTGTCTGTGAAGCAGAGAGTCTTT 352

QY 241 GAAAAAGCTTTTAAAGACTTCTTCCATCACTACTGAGAGAGTGAAGAG 300

DB 353 GAAAAAGCTTTTAAAGACTTCTTCCATCACTACTGAGAGAGTGAAGAG 412

QY 301 ATTATGAGCTAGCTGTAACCGGAGGTAACAGAGATGAGATTATCTTAATAGTACC 360



```
Db 413 ATTTAGAGCTACGTGTAAACCGAGCTACCCAGAGATGATTTTATCTCTAATAGTACC 472
Qy 361 CGCTTATCTTTGAGACCAAGAAATATCTGTCTTCAATCAAAACAGACAGAGGCTTAATAC 420
Db 473 CGCTTATCTTTGAGACCAAGAAATATCTGTCTTCAATCAAAACAGAGGCTTAATAC 532
Qy 421 AAGCCAAAGCAAGAGTGAAGTTTGCATTTGTTACACTCTTCTCAGATTTTAAGCCTTAC 480
Db 533 AAGCCAAAGCAAGAGTGAAGTTTGCATTTGTTACACTCTTCTCAGATTTTAAGCCTTAC 592
Qy 481 AAAACCTCTTTAAACATTTCTCAATTAAAGACCCCAATTCAAATTTGATCCAAGTGTG 540
Db 593 AAAACCTCTTTAAACATTTCTCAATTAAAGACCCCAATTCAAATTTGATCCAAGTGTG 652
Qy 541 TCACAAACAAGTATCTTGGAGTCATTTCCAAAACCTTTACAGTATCTTCCATCCAAAT 600
Db 653 TCACAAACAAGTATCTTGGAGTCATTTCCAAAACCTTTACAGTATCTTCCATCCAAAT 712
Qy 601 CTGTGACTGTCTATTCAGTTCAAGTTCAAGTGAATGACACAGACATATTATCAATCATTTG 660
Db 713 CTGTGAGACTGTCTATTCAGTTCAAGTTCAAGTGAATGACACAGACATATTATCAATCATTTG 772
Qy 661 GTTTGAGATATGTATTTACCAAAATTTGAAGTACTTTGACAGACACATTAATTTGTTCT 720
Db 773 GTTTGAGATATGTATTTACCAAAATTTGAAGTACTTTGACAGACACATTAATTTGTTCT 832
Qy 721 ATGAATTCAGATTTAAATGATGACATACGGGCAAGATTAATCATATGGAAGCCAGTG 780
Db 833 ATGAATTCAGATTTAAATGATGACATACGGGCAAGATTAATCATATGGAAGCCAGTG 892
Qy 781 AAAGAGACGTACGCTTACATTTTACCTTTATCCTTTGGGGAAGAAGAAATATTT 840
Db 893 AAAGAGAGCTACGCTTACATTTTACCTTTATCCTTTGGGGAAGAAGAAATATTT 952
Qy 841 ACAAACAATTTAAGTAAATGATCTGCAAACTTCTCTTTTAATGATGAAGATGAA 900
Db 953 ACAAACAATTTAAGTAAATGATCTGCAAACTTCTCTTTTAATGATGAAGATGAA 1012
Qy 901 AATGTATGATTTCTTCAATGAGACTTTGTAATCTGTAATCTTCCCTGAGACA 960
Db 1013 AATGTATGATTTCTTCAATGAGACTTTGTAATCTGTAATCTTCCCTGAGACA 1072
Qy 961 GTAGAAATTTTAAACCAAGTGAACAGATCAGTTACAGTATTTCAAGAATGTAGACT 1020
Db 1073 GTAGAAATTTTAAACCAAGTGAACAGATCAGTTACAGTATTTCAAGAATGTAGACT 1132
Qy 1021 AATGTGTCTTCAAGACATGATTAATCATTTGAGTTTGTATTATACTACTGTG 1080
Db 1133 AATGTGTCTTCAAGACATGATTAATCATTTGAGTTTGTATTATACTACTGTG 1192
Qy 1081 AAGGCACTCTCAACTTCAACGACCTGGAAGTAACTGCTGATGAGCAACCACTG 1140
Db 1193 AAGGCACTCTCAACTTCAACGACCTGGAAGTAACTGCTGATGAGCAACCACTG 1252
Qy 1141 ACTCTTGAAGAAAGAAATATATGTATGATTAAGTATTAAGTACAAGAGAACTATAGTG 1200
Db 1253 ACTCTTGAAGAAAGAAATATATGTATGATTAAGTATTAAGTACAAGAGAACTATAGTG 1312
Qy 1201 TACTGAGGAGATCTTAACAGTGAATTCAGAAAATGGAAGCTGTTCAAGAAAATTAATAT 1260
Db 1313 TACTGAGGAGATCTTAACAGTGAATTCAGAAAATGGAAGCTGTTCAAGAAAATTAATAT 1372
Qy 1261 ACTGTCCCCCAAGTGAACCTTTTAAGTTGAATTCCTGAGAGATTCAGTGAAG 1320
Db 1373 ACTGTCCCCCAAGTGAACCTTTTAAGTTGAATTCCTGAGAGATTCAGTGAAG 1432
Qy 1321 CTACAGTTGAAGGCTATTTCTGTGTAGTAAAGATGAGAGTGCAGTTGATGTCTGTT 1380
Db 1433 CTACAGTTGAAGGCTATTTCTGTGTAGTAAAGATGAGAGTGCAGTTGATGTCTGTT 1492
Qy 1381 AAGTCTCTAGTAAAGCATATCATCAATTAAGCAAGATGAATTAAGTGAAGTGGAG 1440
Db 1492 AAGTCTCTAGTAAAGCATATCATCAATTAAGCAAGATGAATTAAGTGAAGTGGAG 1552
Qy 1441 TGCCCTTTTGAAGTTGGTGTAGTGGCAACAAACGATTTGAAGAGTATAGTATAGTGA 1500
Db 1553 TGCCCTTTTGAAGTTGGTGTAGTGGCAACAAACGATTTGAAGAGTATAGTATAGTGA 1612
Qy 1501 GTATCAGAGGACAGTTGGTGGCTGTAGGAAAACAAAATTCACAAATGTTCTTTTAACA 1560
Db 1613 GTATCAGAGGACAGTTGGTGGCTGTAGGAAAACAAAATTCACAAATGTTCTTTTAACA 1672
Qy 1561 CCAGAAAATTTCTGGAACCTCCAAAACCTGTGTAAATTTGTATTAATTTGAAGATGATGG 1620
Db 1673 CCAGAAAATTTCTGGAACCTCCAAAACCTGTGTAAATTTGTATTAATTTGAAGATGATGG 1732
Qy 1621 GAATTTAATAGTATGTTCTTAAATAATCTGTGACGTGTTGTTTAAATAATTAAGTAAAG 1680
Db 1733 GAATTTAATAGTATGTTCTTAAATAATCTGTGACGTGTTGTTTAAATAATTAAGTAAAG 1792
Qy 1681 CTATATTTGAGTAAAGTAAAGCTGGAACCATCTGAGAAAGTCTCTTTAGATCTGTG 1740
Db 1793 CTATATTTGAGTAAAGTAAAGCTGGAACCATCTGAGAAAGTCTCTTTAGATCTGTG 1852
Qy 1741 ACACAGCTGACTCATATGTTGGATTTGACTGTGCAAAAAGTGTGAATCTGTGAAT 1800
Db 1853 ACACAGCTGACTCATATGTTGGATTTGACTGTGCAAAAAGTGTGAATCTGTGAAT 1912
Qy 1801 GCCTTAATGATATTAACAATGGAATAATGTCATGAGTGTGAACCTTTATTAACACAGA 1860
Db 1913 GCCTTAATGATATTAACAATGGAATAATGTCATGAGTGTGAACCTTTATTAACACAGA 1972
Qy 1861 TATTAATTTAGGATTTGATGAATTTCTTTGACGCTTTACAGATGTGCACTGAGGTA 1920
Db 1973 TATTAATTTAGGATTTGATGAATTTCTTTGACGCTTTACAGATGTGCACTGAGGTA 2032
Qy 1921 TTGACAGATGCAAACTTCAAGAGATTAATTTGATGTTTATGATCAATGACAGATAT 1980
Db 2033 TTGACAGATGCAAACTTCAAGAGATTAATTTGATGTTTATGATCAATGACAGATAT 2092
Qy 1981 GCTGAGGTTTATGAGAGAAATGAAGACATTTGATGATTTATGATGATGATGATGATG 2040
Db 2093 GCTGAGGTTTATGAGAGAAATGAAGACATTTGATGATTTATGATGATGATGATGATG 2152
Qy 2041 GGTACAGTCCACATGTCGGAAGATTTTCCAGAGACTTGGATTTGGCTAGACACCAAC 2100
Db 2153 GGTACAGTCCACATGTCGGAAGATTTTCCAGAGACTTGGATTTGGCTAGACACCAAC 2212
Qy 2101 ATGGGTTACAGGATTTACCAAGATTTGAAGTAACTGATCTGATCTATCACTTCTGG 2160
Db 2213 ATGGGTTACAGGATTTACCAAGATTTGAAGTAACTGATCTGATCTATCACTTCTGG 2272
Qy 2161 GTGGCTACTGGTTTGTGATCTCTGAGAGACCTGGGTTTGAAGCTTAACAACATCTCCAGTG 2220
Db 2273 GTGGCTACTGGTTTGTGATCTCTGAGAGACCTGGGTTTGAAGCTTAACAACATCTCCAGTG 2332
Qy 2221 GAGCTCCAGGCTTCCAAACATTTTTCATTTTGTGATCTTCCCTACTCTGTTATGAGA 2280
Db 2333 GAGCTCCAGGCTTCCAAACATTTTTCATTTTGTGATCTTCCCTACTCTGTTATGAGA 2392
Qy 2281 GGTGAAGATTTGCTTTGGAATAATCTATATTCATTTTGAAGATGCCATGAGGTT 2340
Db 2393 GGTGAAGATTTGCTTTGGAATAATCTATATTCATTTTGAAGATGCCATGAGGTT 2452
Qy 2341 AAGGTATATCTTGAAGAAAGTGAACAAATTTGATTTCTTAAGACTTCAAGTGAATAAT 2400
Db 2453 AAGGTATATCTTGAAGAAAGTGAACAAATTTGATTTCTTAAGACTTCAAGTGAATAAT 2512
Qy 2401 GCCACAGGCCACAGACAGACCTTCTGTGTCCAGTGAAGATGGGCAACTGTTCTTTT 2460
Db 2513 GCCACAGGCCACAGACAGACCTTCTGTGTCCAGTGAAGATGGGCAACTGTTCTTTT 2572
Qy 2461 CCCATCAGGCCAACAATCTGTGGAGAAATTTCTATCAAGTCAAGCTCTTTCAACCACT 2520
Db 2573 CCCATCAGGCCAACAATCTGTGGAGAAATTTCTATCAAGTCAAGCTCTTTCAACCACT 2632
```

QY 2521 GCTTCGATGCTGTCACCCAGATGATTTTAAAGGCTGAAGAAATAGAAAAATCATAT 2580  
 Db 2633 GCTTCGATGCTGTCACCCAGATGATTTTAAAGGCTGAAGAAATAGAAAAATCATAT 2692  
 QY 2581 TCACATCATCTTATTTAGACTGACGACATATGAGCTACAGAGTACCTGAAAACTTTG 2640  
 Db 2693 TCACATCATCTTATTTAGACTGACGACATATGAGCTACAGAGTACCTGAAAACTTTG 2752  
 QY 2641 AGTTTCATCTTCCCTTAATAGCTGACCTGCGAGTGAAGAAAGTTCCAGATCTGCAATT 2700  
 Db 2753 AGTTTCATCTTCCCTTAATAGCTGACCTGCGAGTGAAGAAAGTTCCAGATCTGCAATT 2812  
 QY 2701 GGAGATGTTCTTGGCTCTTCATCAATGCTTAGCCTCATTTGATTCGAGCTTAAATGCG 2760  
 Db 2813 GGAGATGTTCTTGGCTCTTCATCAATGCTTAGCCTCATTTGATTCGAGCTTAAATGCG 2872  
 QY 2761 TGTGCTGAACAGACATGATATAATTTGCTCCAAATATTTACATTTTGGATTATCTGACT 2820  
 Db 2873 TGTGCTGAACAGACATGATATAATTTGCTCCAAATATTTACATTTTGGATTATCTGACT 2932  
 QY 2821 AAAAAAGAAACAACGACGATTAATTTGAAGAAAGCTTTTCATTTATAGGCAAGCT 2880  
 Db 2933 AAAAAAGAAACAACGACGATTAATTTGAAGAAAGCTTTTCATTTATAGGCAAGCT 2992  
 QY 2881 TACAGAGAGAACTCTCTATCAGAGGAGAGATGCTCTTCACTGCTTTTGGAAATTTAT 2940  
 Db 2993 TACAGAGAGAACTCTCTATCAGAGGAGAGATGCTCTTCACTGCTTTTGGAAATTTAT 3052  
 QY 2941 GACCTTCTTGGAGCACTTGTGTGCTGCTTTGTTTAAAGATGTTTCTTGAAGCCGAT 3000  
 Db 3053 GACCTTCTTGGAGCACTTGTGTGCTGCTTTGTTTAAAGATGTTTCTTGAAGCCGAT 3112  
 QY 3001 CCTTACATAGATTTTGAATGACAAATGCTTTTACACAGAAATACCTTGGCTTAAAGCAAT 3060  
 Db 3113 CCTTACATAGATTTTGAATGACAAATGCTTTTACACAGAAATACCTTGGCTTAAAGCAAT 3172  
 QY 3061 CAGAAATCAACGCTGATTTTGGGATTCAGAGAGAGATTTCACTTGAAGCTTCAAGCT 3120  
 Db 3173 CAGAAATCAACGCTGATTTTGGGATTCAGAGAGAGATTTCACTTGAAGCTTCAAGCT 3232  
 QY 3121 GGCATTAATAAGCTCAGTAACCTTACAGCTTAATTTGTAACCTTCTCTGGGATATAGA 3180  
 Db 3233 GGCATTAATAAGCTCAGTAACCTTACAGCTTAATTTGTAACCTTCTCTGGGATATAGA 3292  
 QY 3181 AAGATACGCTTACATTTGATGTCAGAGATCTTCACTTGGGATCTGATTTCACT 3240  
 Db 3293 AAGATACGCTTACATTTGATGTCAGAGATCTTCACTTGGGATCTGATTTCACT 3352  
 QY 3241 AGAGAAATTTTCAACAATTTACTCTAGCCCTTAACTTATGATTTGTCATGAGTGGG 3300  
 Db 3353 AGAGAAATTTTCAACAATTTACTCTAGCCCTTAACTTATGATTTGTCATGAGTGGG 3412  
 QY 3301 AGTCTTAAGGGAAGAAAGCTTTGATATGCTGAATTTGAGAGAGACAAAGAAAGTGGC 3360  
 Db 3413 AGTCTTAAGGGAAGAAAGCTTTGATATGCTGAATTTGAGAGAGACAAAGAAAGTGGC 3472  
 QY 3361 ATGCAATTTCTGGGTGTCTACAGATCTTCACTTCTGCTTGGGACCCAGCTCTCCCTG 3420  
 Db 3473 ATGCAATTTCTGGGTGTCTACAGATCTTCACTTCTGCTTGGGACCCAGCTCTCCCTG 3532  
 QY 3421 GATATTTGAAGTTGAGCTTATGACCTGCTTCACTTCACTTCAATTTCAAGCTTTGAG 3480  
 Db 3533 GATATTTGAAGTTGAGCTTATGACCTGCTTCACTTCACTTCAATTTCAAGCTTTGAG 3592  
 QY 3481 GGAATCCCAATTAAGAGTGTCTAAGCAGGCAAGAAATAGCTTGGTGTGTTTGCATCT 3540  
 Db 3593 GGAATCCCAATTAAGAGTGTCTAAGCAGGCAAGAAATAGCTTGGTGTGTTTGCATCT 3652  
 QY 3541 ACTCAGGATACCACTGTGGCTTTTAAAGGCTTGTCTGATTTTGAAGCCCTTATGAAATCA 3600  
 Db 3653 ACTCAGGATACCACTGTGGCTTTTAAAGGCTTGTCTGATTTTGAAGCCCTTATGAAATCA 3712

QY 3601 GAAAGCAAAATATCCAGATGACCTGACGCGGCTTAGCTCACAACTCTT----- 3651  
 Db 3713 GAAAGCAAAATATCCAGATGACCTGACGCGGCTTAGCTCACAACTCTTAAAGTTT 3772  
 QY 3652 -----CTTGCTGTGTGACGCCA 3669  
 Db 3773 CTGATTGACACACAAACCGCTTATCTCTTCAACAGACAGAGCTGTGTGTGACGCCA 3832  
 QY 3670 ATGGCAGTTAATATTTTCCGAAATGTTTGAATTTGCTTATTTGTCAGCTCAATGTTGTA 3729  
 Db 3833 ATGGCAGTTAATATTTTCCGAAATGTTTGAATTTGCTTATTTGTCAGCTCAATGTTGTA 3892  
 QY 3730 TATATGTGAAGCTTCTGGGCTTCTTAGAAGACGAAATCTATTCAAAAATCAAGAACCC 3789  
 Db 3893 TATATGTGAAGCTTCTGGGCTTCTTAGAAGACGAAATCTATTCAAAAATCAAGAACCC 3952  
 QY 3790 TTTGATTTAATGTTGTGTAAAGAAATTAAGATGATCTCAATGATGATTTGTAAT 3849  
 Db 3953 TTTGATTTAATGTTGTGTAAAGAAATTAAGATGATCTCAATGATGATTTGTAAT 4012  
 QY 3850 GTGTGTCAAGCTTTTCCGCGCCCGGCTAGGAGTGGCATGCTCTTATGAAATTAACCTTA 3909  
 Db 4013 GTGTGTCAAGCTTTTCCGCGCCCGGCTAGGAGTGGCATGCTCTTATGAAATTAACCTTA 4072  
 QY 3910 TTAAGTGGCTTTATGCTGCTTCAAGAACATTTCTTGAAGCAGACAGTGAAGAAAGTG 3969  
 Db 4073 TTAAGTGGCTTTATGCTGCTTCAAGAACATTTCTTGAAGCAGACAGTGAAGAAAGTG 4132  
 QY 3970 GAATATGATATGGAAGAACTCAACCTTATTTAGATTTCTGTAATGAAACCCAGTTTGT 4029  
 Db 4133 GAATATGATATGGAAGAACTCAACCTTATTTAGATTTCTGTAATGAAACCCAGTTTGT 4192  
 QY 4030 GTTAATATCTCTGCTGAGAACTTTAAAGTTTCAATCCCAAGATGCTTCACTGCTC 4089  
 Db 4193 GTTAATATCTCTGCTGAGAACTTTAAAGTTTCAATCCCAAGATGCTTCACTGCTC 4252  
 QY 4090 ATATGATATTAATGAGCCAGAGAGACAGCGGTGAGAAATTAACCTTGAAGTGAAG 4149  
 Db 4253 ATATGATATTAATGAGCCAGAGAGACAGCGGTGAGAAATTAACCTTGAAGTGAAG 4312  
 QY 4150 CTGTCTCTCTGTGACCTTTGCAATGATGTCAGAGGCTGCGCTCTTGTGAGATGAGCT 4209  
 Db 4313 CTGTCTCTCTGTGACCTTTGCAATGATGTCAGAGGCTGCGCTCTTGTGAGATGAGCT 4372  
 QY 4210 TCAGGCTCCCATGATCACTTTCAGTCAATTTTATTTTGTTCAGAGCTTCTGACTTT 4269  
 Db 4373 TCAGGCTCCCATGATCACTTTCAGTCAATTTTATTTTGTTCAGAGCTTCTGACTTT 4432  
 QY 4270 ATGAACTTTGGCTGTGA 4287  
 Db 4433 ATGAACTTTGGCTGTGA 4450

RESULT 3  
 AX534938  
 LOCUS 5895 bp DNA linear PAT 22-NOV-2002  
 DEFINITION Sequence 5 from Patent WO02070696.  
 ACCESSION AX534938  
 VERSION AX534938.1 GI:25261482  
 KEYWORDS  
 SOURCE  
 ORGANISM  
 Homo sapiens (human)  
 Homo sapiens  
 Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.  
 REFERENCE  
 1 Schuh, A. and Sutherland, R.D.  
 CD109 nucleic acid molecules polypeptides and methods of use  
 Patent: WO 02070696-A 5 12-SEP-2002;  
 Schuh, Andrew Sutherland, Robert D. (CA)  
 Location/Qualifiers  
 1..5895  
 /organism="Homo sapiens"  
 /mol\_type="genomic DNA"

Same as 1

CDS

/db\_xref="taxon:9606"  
113..4450  
/note="unnamed protein product"  
/codon\_start=1  
/protein\_id="CAD57233.1"  
/db\_xref="GI:25261483"  
/translation="MQGPPPLTRAAHLCCVTAALAAVAPSPRLVTPAGIIRGNVTT  
GVELLEHCPSPVTVAAELKLTASNLTVSLAEAVPEKSFETLLPSLPLNSADELY  
ELRVGTQDEILFNSNTRLSFEETKIVSYFIDTKALYPRKQEVRIYTLFSDPEKY  
KSLNLIIDKPSNLIOQMLSOOSDGLSVKTFOLSHSHIIDGMSIQOVNOQTYOS  
FOVSEYLPKPEVTLQOTPLVCGSMNSKHNLGTLTAKTYKPKVGVTLTFLPSWTK  
KNITTEKINGSANSPFDEEMKYMDSNGLSEYDLSSEGPBEILITTTSTMG  
SNRVSTNFFKQHDITIEFFDTTVLKSPLNTAVKTRAGNOULTEERNNAVIT  
VTQRYNTEYSSNSGNQCEAVQKINYPQSGTFKLEPILLESSELOKNGEYFV  
KSSMAVHSLFKSPSKTYIQKTRDENIKVSPFELVANSKELKEISLVYVARGDLVA  
VGNKSTMFSLTPENSPTKACIVIVYIIDDEGILSDVLTIVOLFKKILKLYSKY  
KAPSEKSLRISTVOPDSIVGIIVAVDKINVMANSDNITMENVAHELKLYMGVYLG  
MEMNSPFAVQEGGLVLTPLANTKDYIIDVYDNAAEAPFMEENSHYDIDFSLGS  
SPHVRKHPPEIWMIDTNGKRIYQEFVTVDSITSWATGFSVISEDGLATTPPV  
ELQAFQPEFIPLNLPYSVIRGEFALEITIENYLDKATVKKVIEKSDSELOKNGEYFV  
INATGQOTLVPSEDAVLPEPIRTHLGEIPITVLTALSPASDAVTOMILVKAEG  
EKSYQSILDLIDNRLOSTLKTLSFPSPNTVTSERVOITAIQVIGPSINGLASL  
IMPYGCGEOMINFPANVIYIIDYLTKKQOLTDNKEKALSPMGQYKQRELLXQREDS  
SPSAGNTVPSGSTWLSAFVLCFPLEADPYIDIDQVTLRHTTTLWLGKNGEYFV  
GVHISELQGNKSPVTLTAIVLTSILGRKQPNIDVGESIHFLSESESRISDNVT  
LALITVYALSVSQPKAKALNMLTWRAEQEGMOPVSSSESLSDSMOPRSIDIEVAA  
VALSHFLQFOTSEGIPIRMWLSRONSLSEGFASVODTVALKALSEFALANTERTIN  
IOVTVGPSPPKVEKIDITHRLLQTELAIVODMAVNIANGGFAICOLNVVN  
VWASSRRRSRISIOEAFDLVAVKENDDLNHYDINACTSPSGRSGMALMEVNL  
LSGRNVPSEALISLSETVKVEYDHEKLNLYLDSVNETQCVNIPARNKVSNTDAS  
VSIIVDEYERQAVRSYNSVSLSCDLSQVQGRPCEDASGSHHSVIFIFCFK  
LYFPEMLWLM"

BASE COUNT 1709 a 1173 c 1241 g 1772 c

ORIGIN

Query Match 98.5%; Score 4224.4; DB 6; Length 5895;  
Best Local Similarity 98.8%; Pred. No. 0;  
Matches 4286; Conservative 0; Mismatches 1; Indels 51; Gaps 1;

1 ATGCAAGGCGCCACCGCTCTGACCGCGCGCCACCTCTCGGTGCGTGAACCGCGCGCGC 60  
113 ATGCAAGGCGCCACCGCTCTGACCGCGCGCCACCTCTCTGCGTGAACCGCGCGCGC 172  
61 GCCGTGCTCCCGGCGCTCGTCTGTGACAGCCCGAGGATCATCAGCCCGGAGGA 120  
173 GCGGTGCTCCCGGCGCTCGTCTGTGACAGCCCGAGGATCATCAGCCCGGAGGA 232  
121 AATGTACTATTGGGGTGAAGCTTCTGAAACCTGECCTTCAAGGTGACTGTGAAGGCG 180  
233 AATGTACTATTGGGGTGAAGCTTCTGAAACCTGECCTTCAAGGTGACTGTGAAGGCG 292  
181 GAGCTGCTAAGACAGATCAAACTCATCTGTCTGTCTGGAAGCAGAGAGCTTT 240  
293 GAGCTGCTAAGACAGATCAAACTCATCTGTCTGTCTGGAAGCAGAGAGCTTT 352  
241 GAAAAAGCTCTTTTAAAGACTTACTCTTCATCTACTCTCTGAAAGCTGGAAGTGA 300  
353 GAAAAAGCTCTTTTAAAGACTTACTCTTCATCTACTCTCTGAAAGCTGGAAGTGA 412  
301 ATTATGAGCTACGTGTAAACCGGACGTACCCAGAGATGAGATTTTATCTCTAATAGTACC 360  
413 ATTATGAGCTACGTGTAAACCGGACGTACCCAGAGATGAGATTTTATCTCTAATAGTACC 472  
361 CGCTATCATTTTGAAGCAGAGAAATATCTGTCTTCATCAACAGAGAAAGCTTATAC 420  
473 CGCTATCATTTTGAAGCAGAGAAATATCTGTCTTCATCAACAGAGAAAGCTTATAC 532  
421 AAGCCAAAGCAAGATGAGATTCGATTTGATCACTCTCTCAGATTTTAAAGCTTAC 480  
533 AAGCCAAAGCAAGATGAGATTCGATTTGATTCATCTCTCAGATTTTAAAGCTTAC 592  
481 AAAAAGCTTTTAAAGATTTCTCATTTAAGACCCCAAAATCAAAATTTGATTCAGAGTGGTG 540

593 AAAAAGCTTTTAAAGATTTCTCATTTAAGACCCCAAAATCAAAATTTGATTCAGAGTGGTG 652  
541 TCACAAAGCAAGATCTTGAAGTCAATTTCAAAATCTTTGAGCTATCTTCCATCCAAATA 600  
653 TCACAAAGCAAGATCTTGAAGTCAATTTCAAAATCTTTGAGCTATCTTCCATCCAAATA 712  
601 CTTGATGACGTGCTTATTCAGAGTCAAGTGAATGACAGACATATTTATCAATCATTTGAG 660  
713 CTTGATGACGTGCTTATTCAGAGTCAAGTGAATGACAGACATATTTATCAATCATTTGAG 772  
661 GTTTAGAAATATATTTTACCAAAATTTGAAGTCAATTTGACAGACATATTTATTTGTTCT 720  
773 GTTTAGAAATATATTTTACCAAAATTTGAAGTCAATTTGACAGACATATTTATTTGTTCT 832  
721 ATGAATTTCAAGCATTTTAAATGATGATCAAGGCAAGTATACATATGGAAGGACAGTG 780  
833 ATGAATTTCAAGCATTTTAAATGATGATCAAGGCAAGTATACATATGGAAGGACAGTG 892  
781 AAAGAGACGTAAAGCTTACATTTTACCTTTATCTTTTGGGAAAGAAATAATTT 840  
893 AAAGAGACGTAAAGCTTACATTTTACCTTTATCTTTTGGGAAAGAAATAATTT 952  
841 AAAAAAATTTAAGATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 900  
953 AAAAAAATTTAAGATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1012  
901 AATGTAATGATTTCTTCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 960  
1013 AATGTAATGATTTCTTCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1072  
961 GTAGAAATTTTAAACCAAGTGAACAGATCAAGTATGATGATGATGATGATGATGATGATGAT 1020  
1073 GTAGAAATTTTAAACCAAGTGAACAGATCAAGTATGATGATGATGATGATGATGATGATGAT 1132  
1021 AATGTTCTTCAAGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1080  
1133 AATGTTCTTCAAGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1192  
1081 AAGCATCTTCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1140  
1193 AAGCATCTTCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1252  
1141 ACTCTTGAAGAAAGAAATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1200  
1253 ACTCTTGAAGAAAGAAATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1312  
1201 TACTGAGCGGATCTTAACAGTGAAGAAATGAGAAATGAGAAATGAGAAATGAGAAATGAGAAAT 1260  
1313 TACTGAGCGGATCTTAACAGTGAAGAAATGAGAAATGAGAAATGAGAAATGAGAAATGAGAAAT 1372  
1261 ACTGTCCTTCAAGTGAAGAAATGAGAAATGAGAAATGAGAAATGAGAAATGAGAAATGAGAAAT 1320  
1373 ACTGTCCTTCAAGTGAAGAAATGAGAAATGAGAAATGAGAAATGAGAAATGAGAAATGAGAAAT 1432  
1321 CTACAGTGAAGGCTTATTTCTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1380  
1433 CTACAGTGAAGGCTTATTTCTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1492  
1381 AAGTCTCTTCAAGTGAAGAAATGAGAAATGAGAAATGAGAAATGAGAAATGAGAAATGAGAAAT 1440  
1493 AAGTCTCTTCAAGTGAAGAAATGAGAAATGAGAAATGAGAAATGAGAAATGAGAAATGAGAAAT 1552  
1441 TCGCTTTTGAAGTGGTGGTGAAGGCAAAAGATTTGAAGGATTTGAAGGATTTGAAGGATTTGAAG 1500  
1553 TCGCTTTTGAAGTGGTGGTGAAGGCAAAAGATTTGAAGGATTTGAAGGATTTGAAGGATTTGAAG 1612  
1501 GTATCAGAGGACAGTGGTGGTGAAGGCAAAAGATTTGAAGGATTTGAAGGATTTGAAGGATTTGAAG 1560  
1613 GTATCAGAGGACAGTGGTGGTGAAGGCAAAAGATTTGAAGGATTTGAAGGATTTGAAGGATTTGAAG 1672  
1561 CCGAGAAATTTTGAAGTCCAAAGCTGTGTAAATGTGTAAATTTGAAGGATTTGAAGGATTTGAAG 1620  
1673 CCGAGAAATTTTGAAGTCCAAAGCTGTGTAAATGTGTAAATTTGAAGGATTTGAAGGATTTGAAG 1732

QY 1621 GAAATTATAGTATGTTCTAAAAATTCCTGTTCACTGTTGTTTTTAAAAATPAAGATAAG 1680  
Db 1723 GAAATTATAGTATGTTCTAAAAATTCCTGTTCACTGTTGTTTTTAAAAATPAAGATAAG 1792  
QY 1681 CTATATTTGAGTAAAGTAAAGCTGACCATCTGAGAAAGTCTCTCTTAGATCTCTGTG 1740  
Db 1793 CTATATTTGAGTAAAGTAAAGCTGACCATCTGAGAAAGTCTCTCTTAGATCTCTGTG 1852  
QY 1741 ACACAGCTGACTCATAGTTGGGATTTGAGTGTGACAAAGAGTGAATCTGATGAT 1800  
Db 1853 ACACAGCTGACTCATAGTTGGGATTTGAGTGTGACAAAGAGTGAATCTGATGAT 1912  
QY 1801 GCTCTTAATGATATTACAAATGAAAATGTGTCATGAGTTGGAACCTTTATPACACAGA 1860  
Db 1913 GCTCTTAATGATATTACAAATGAAAATGTGTCATGAGTTGGAACCTTTATPACACAGA 1972  
QY 1861 TATATTTAGGCAATGTTATGATATTTCTTTGACAGTCTTTCAGAAATGGAATCTGGGTA 1920  
Db 1973 TATATTTAGGCAATGTTATGATATTTCTTTGACAGTCTTTCAGAAATGGAATCTGGGTA 2032  
QY 1921 TTGACAGATGCAAACTCAGCAAGGATTAATATGATGTTGTTTATGACATGACGAATAT 1980  
Db 2033 TTGACAGATGCAAACTCAGCAAGGATTAATATGATGTTGTTTATGACATGACGAATAT 2092  
QY 1981 GCTGAGAGTTTATGAGAAAATGMAAGACATATTTGATATATTCATGACTTTTCTTTG 2040  
Db 2093 GCTGAGAGTTTATGAGAAAATGMAAGACATATTTGATATATTCATGACTTTTCTTTG 2152  
QY 2041 GGTAGCAATCCACATGTCGAAAGCATTTTCCAGAGCTTGGATTTGGCTGACACCAAC 2100  
Db 2153 GGTAGCAATCCACATGTCGAAAGCATTTTCCAGAGCTTGGATTTGGCTGACACCAAC 2212  
QY 2101 ATGGATTAACAGATTTTACCAAGAAATTTGAGTAATCTGATCTGATCACTTTGG 2160  
Db 2213 ATGGATTAACAGATTTTACCAAGAAATTTGAGTAATCTGATCTGATCACTTTGG 2272  
QY 2161 GTGGCTACTGTTTGTGATCTCTGAGAGCTGGGCTTGGACTPACAACTACTCCAGTG 2220  
Db 2273 GTGGCTACTGTTTGTGATCTCTGAGAGCTGGGCTTGGACTPACAACTACTCCAGTG 2332  
QY 2221 GAGCTCCAGGCTTCCAAACATTTTCACTTTTGTGAATCTTCCCTACTCTGTTATCA 2280  
Db 2333 GAGCTCCAGGCTTCCAAACATTTTCACTTTTGTGAATCTTCCCTACTCTGTTATCA 2392  
QY 2281 GGTAAAGATTTGCTTGGAAATACTATATTCATTAATTTGAAGAGTCCAGTGGGT 2340  
Db 2393 GGTAAAGATTTGCTTGGAAATACTATATTCATTAATTTGAAGAGTCCAGTGGGT 2452  
QY 2341 AAGGTAATCATTTGAAAAAGTACAAATTTGATATTTCAATGACTTCAAGTAATAAT 2400  
Db 2453 AAGGTAATCATTTGAAAAAGTACAAATTTGATATTTCAATGACTTCAAGTAATAAT 2512  
QY 2401 GCCACAGGCAACAGACAGACCTTCTGTGTTCCAGTGAAGATGGGCAACTGTTCTTTT 2460  
Db 2513 GCCACAGGCAACAGACAGACCTTCTGTGTTCCAGTGAAGATGGGCAACTGTTCTTTT 2572  
QY 2461 CCCATCAGGCCAACAACATCTGGAGAAATTTCTATACAGTCAAGCTCTTTCACCACT 2520  
Db 2573 CCCATCAGGCCAACAACATCTGGAGAAATTTCTATACAGTCAAGCTCTTTCACCACT 2632  
QY 2521 GCTTCTGATGCTGTACCCAGATGATTTTATGAAAGCTGAGAAATGAAAAATCATAT 2580  
Db 2633 GCTTCTGATGCTGTACCCAGATGATTTTATGAAAGCTGAGAAATGAAAAATCATAT 2692  
QY 2581 TCACAAATCCATCTTATTAAGCTTGACTGACAAATAGGCTACAGAGTACCTGAAAACTTTG 2640  
Db 2693 TCACAAATCCATCTTATTAAGCTTGACTGACAAATAGGCTACAGAGTACCTGAAAACTTTG 2752  
QY 2641 AGTTTCATTTCTCTCAATATACAGTGACTGGCACTGAAAAGTTGAGATTCAGTGCATAT 2700  
Db 2753 AGTTTCATTTCTCTCAATATACAGTGACTGGCACTGAAAAGTTGAGATTCAGTGCATAT 2812

QY 2701 GAGATGTTCTTGCTCTTCATCATATGCTTACCTTACCTCATTTGATTCGAGTGCCTTATGCG 2760  
Db 2813 GAGATGTTCTTGCTCTTCATCATATGCTTACCTTACCTCATTTGATTCGAGTGCCTTATGCG 2872  
QY 2761 TGTGTAACAGAAACATGATTAATTTGCTCCAAATTTTACATTTGATGATTCGACT 2820  
Db 2873 TGTGTAACAGAAACATGATTAATTTGCTCCAAATTTTACATTTGATGATTCGACT 2932  
QY 2821 AAAAAGAAACAATGACAGATTAATTTGAAAAGAAAAGCTCTTCAATTAATGAGCAAGT 2880  
Db 2933 AAAAAGAAACAATGACAGATTAATTTGAAAAGAAAAGCTCTTCAATTAATGAGCAAGT 2992  
QY 2881 TACCAGAGAACTTCTCTATACAGGGAAGATGCTCTTCACTGCTTTTGGGAATTA 2940  
Db 2993 TACCAGAGAACTTCTCTATACAGGGAAGATGCTCTTCACTGCTTTTGGGAATTA 3052  
QY 2941 GACCTTCTGGAGCACTTGTTGTACGCTTTGTTTAAAGATTTCTTGAAGCCGAT 3000  
Db 3053 GACCTTCTGGAGCACTTGTTGTACGCTTTGTTTAAAGATTTCTTGAAGCCGAT 3112  
QY 3001 CCTTACATAGATATTTGATCAGAAATGTTTACAGAAACATACACTTGGCTTAAAGACAT 3060  
Db 3113 CCTTACATAGATATTTGATCAGAAATGTTTACAGAAACATACACTTGGCTTAAAGACAT 3172  
QY 3061 CAGAAATCCAAAGGTGAATTTGGGATCCAGAAAGAGTGAATCAATGAGCTTCAAGT 3120  
Db 3173 CAGAAATCCAAAGGTGAATTTGGGATCCAGAAAGAGTGAATCAATGAGCTTCAAGT 3232  
QY 3121 GGCATTAAGTCCAGTAACATTAACAGCTTATATTTGTAATCTTCTCCCTGGATATAGA 3180  
Db 3233 GGCATTAAGTCCAGTAACATTAACAGCTTATATTTGTAATCTTCTCCCTGGATATAGA 3292  
QY 3181 AAGTACAGCTTAACATTTGATGCAAGAGTCTATCCATTTTGGAGTGTGAATTCAGT 3240  
Db 3293 AAGTACAGCTTAACATTTGATGCAAGAGTCTATCCATTTTGGAGTGTGAATTCAGT 3352  
QY 3241 AGAGAAATTTCAACAATTAATTAATCTTACCTTATTAATCTTATGATGTCATGATGGG 3300  
Db 3353 AGAGAAATTTCAACAATTAATTAATCTTACCTTATTAATCTTATGATGTCATGATGGG 3412  
QY 3301 AGTCTTAACAGCAAGAGGCTTGAATATGCTGACTTGGAGACAGAAAGAGTGGC 3360  
Db 3413 AGTCTTAACAGCAAGAGGCTTGAATATGCTGACTTGGAGACAGAAAGAGTGGC 3472  
QY 3361 ATGCAATTTGGGATGATCAGAGTCCAAACTTTGCACTCCCTGGAGCAGCAGCTCCCTG 3420  
Db 3473 ATGCAATTTGGGATGATCAGAGTCCAAACTTTGCACTCCCTGGAGCAGCAGCTCCCTG 3532  
QY 3421 GATATTAAGTTCAGAGCTTAATGACATGCTCTCACTTCTTACAAATTTCAAGCTTCTGAG 3480  
Db 3533 GATATTAAGTTCAGAGCTTAATGACATGCTCTCTCACTTCTTACAAATTTCAAGCTTCTGAG 3592  
QY 3481 GGAATCCCAATTAATGAGTGGCTTAAGCAGGCAAGAAATAGCTTGGGTGTTTGCATCT 3540  
Db 3593 GGAATCCCAATTAATGAGTGGCTTAAGCAGGCAAGAAATAGCTTGGGTGTTTGCATCT 3652  
QY 3541 ACTCAGATACACTGAGCTTAAAGGCTCTGCTGCAATTTGCAAGCCCTTAATGAATACA 3600  
Db 3653 ACTCAGATACACTGAGCTTAAAGGCTCTGCTGCAATTTGCAAGCCCTTAATGAATACA 3712  
QY 3601 GAAAGCAAAATATCCAAAGTGAACGCTGACGAGGCTTAAGCTCAACAACTCT----- 3651  
Db 3713 GAAAGCAAAATATCCAAAGTGAACGCTGACGAGGCTTAAGCTCAACAACTCTGAAAGTTT 3772  
QY 3652 -----CTTGTGTGTGTAACGCA 3669  
Db 3773 CTGATGACACACAAACGCTTAATCTTCTTACAGACAGAGCTTCTGTGTGTAACGCA 3832  
QY 3670 ATGGAGTTTAATTTTCCGCAAAATGTTTGAATTTGATTTGATGAGCTCAATGTTGTA 3729  
Db 3833 ATGGAGTTTAATTTTCCGCAAAATGTTTGAATTTGATTTGATGAGCTCAATGTTGTA 3892  
QY 3730 TATATGTAAGGCTTCTGGGCTTCTAGAAAGCAAGATCTATCCAAATCAAGAGCC 3789

```

|||||
Db      3893 TATAAGTAAAGGCTTCTGCTCTCTTAAAGCGAAGATCTATCAAAATCAAGAGGC 3952
|||
Qy      3790 TTTGATTTAGATGTTGCTGTAAGAAATTAAGATGATCTCAATCATGATGATTTGAAT 3849
|||
Db      3953 TTTGATTTAGATGTTGCTGTAAGAAATTAAGATGATCTCAATCATGATGATTTGAAT 4012
|||
Qy      3850 GTGTGTACAAAGCTTTTCGGGCCCGGAGTAGAGTGGAGCTTTTATGGAAGTTAACTTA 3909
|||
Db      4013 GTGTGTACAAAGCTTTTCGGGCCCGGAGTAGAGTGGAGCTTTTATGGAAGTTAACTTA 4072
|||
Qy      3910 TTTAGTGGCTTTATGCTGCTTCAAGACATTTCTCTGAGCGAGACAGTGAAGAAATG 3969
|||
Db      4073 TTTAGTGGCTTTATGCTGCTTCAAGACATTTCTCTGAGCGAGACAGTGAAGAAATG 4132
|||
Qy      3970 GAATATGATCATGAGAAACCTCAACCTATTTAGATTCTGTAAATGAATCAACGATTTTCT 4029
|||
Db      4133 GAATATGATCATGAGAAACCTCAACCTATTTAGATTCTGTAAATGAATCAACGATTTTCT 4192
|||
Qy      4030 GTTAATATTTCTGCTGTGAGAAACTTTAAAGTTCAATACCAAGATGCTTCAAGTCTC 4089
|||
Db      4193 GTTAATATTTCTGCTGTGAGAAACTTTAAAGTTCAATACCAAGATGCTTCAAGTCTC 4252
|||
Qy      4090 ATATGATGATTTACTATGAGCCAAAGAGACAGCGCGTGAAGATTACACTCTGAAGTGAAG 4149
|||
Db      4253 ATATGATGATTTACTATGAGCCAAAGAGACAGCGCGTGAAGATTACACTCTGAAGTGAAG 4312
|||
Qy      4150 CTGTCTCTGCTGAGACCTTTGAGTAGATGTCAGAGGCGTGCCTCTGTGAGATGAGT 4209
|||
Db      4313 CTGTCTCTGCTGAGACCTTTGAGTAGATGTCAGAGGCGTGCCTCTGTGAGATGAGT 4372
|||
Qy      4210 TCAGGCTCCCATCATCATCTTTCAGTCAATTTTATTTCTGTTCAAGCTTCTGACTT 4269
|||
Db      4373 TCAGGCTCCCATCATCATCTTTCAGTCAATTTTATTTCTGTTCAAGCTTCTGACTT 4432
|||
Qy      4270 ATGGAACCTTGGCTGTGA 4287
|||
Db      4433 ATGGAACCTTGGCTGTGA 4450
|||

RESULT 4
AX534936          4761 bp      DNA      linear      PAT 22-NOV-2002
LOCUS             Sequence 3 from Patent WO02070696.
DEFINITION        AX534936
ACCESSION         AX534936
VERSION           AX534936.1 GI:25261478
KEYWORDS
SOURCE            Homo sapiens (human)
ORGANISM          Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.

REFERENCE
1 Schu, A. and Sutherland, R.D.
AUTHORS           Cdl09 nucleic acid molecules polypeptides and methods of use
TITLE             Patent: WO 02070696-A 3 12-SEP-2002;
JOURNAL           Schu, Andre (CA) ; Sutherland, Robert D. (CA)
FEATURES
source            Location/Qualifiers
1..4761
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
113..4450
/note="unnamed protein product"
/codon_start=1
/protein_id="CAD57232.1"
/db_xref="GI:25261479"
/translation="MQGPPLTAALHLCVCTAALAVAPGRFLVTAAGIIRPGANVTI
GVELHCPSOVTVKAEILKTAENLTVSLVLAAGVEKESFETLTPLSLNSADEIT
ELRVGRDDELFSNTRLSPEYKRSISVFIOTDKALYKPKQVRYLTLSDEKPY
KISNLILIDPKNLILQOMLSQSDIGVSKTFOLSSHILDSMTIQQVNDQTYQS
FOVSEYLVKFEVTLITPLVYCSMNSHGLGTTTAKYATKPKVGDVTLTFLTESFWK
KNITKTFKINSANFSFDEBKVMDSNGLSEYLDLSFGPEVILTLTESVTGI
SNVSTNVFEKHDIYIEFPDYTVLKPISLFTATVKTVRADNQLTEERRNNVIT

```

Y and as  
Res. 1

```

VTORNYTEWGSNSGKMEAVOKINITYVROSGTFKIEPILEDSSSELQKAYFLGS
KSMVAHSLFKSPKTYIOLTRDENIIVGSPFELVYVGNRLKELSMVAVROSLVA
VKGONWTFSPKACVIVYVIEDGELISDLKXIPVOLVEXKILKWSKY
KABSEKSLISTOPDSIVGIYAVKSVNLVMAINSINTMENVHLELVNTGYLGS
MFMNSFAYFOECGLMTLDNLTNDYIDVGVDAEYARFENEBNGLVLDHDSLGS
SPHAKHPETWIMLDTNWGSRIIOEFVTVPDSTVAVGPIVLEBGLVLDHDSLGS
ELOAFOPPIFLNPSVYIRGEFALERTINYLKADAEVVIIEKSPKFLIMTNE
INAGHOOTLLVPEDGATVLEPIRPHLGEIPIVTVLSPASDAVOMILVKAEG
EKSYSGSILDLTDNRLOSTLKTLSFSPRPNTVGSERVOTIAGDVGPSINASTI
IRMPYCGSEQKMINIPAPITYILDTITKKKQITDNLKELSPMQGYRELLYQREDG
SPSAGNDPESGSTMFAVLCRUEADPVIDIDONVLRITYMLKHOXSGEFTWD
GRVHSELQGNKSPVLTALYVSLGVRKYOPNIDVQESIHFLSEFSRGISDNVY
LALITVALSVSGSPKAKMLMLTMRALGEGMOFWVSESKLSPOPSRSLDVEAA
YALISHFLOFOTSGEIPMRWLSROBNLSGEGFASODPTVALKALSEPAALMTERTN
IOVTVTGSSPSVPKFLIDTNRLILOTAELAVOPMNVINSANGFALIOANVYN
VKAGSSRRRSRSIONOEAFDIDVAVKENDLNDVNDVNCISFSGRSGMALEVNIL
LSGFVPEBAISLSEYAKVYDHKMLYLDSEVNEFQVNIIPAVRNFKVSNTODAS
VSIYDYVEPRQAVRSYNSVEKLSCDLCSVDQCRPCEDGASGSHHSVIFLFCFK
LLYFPEMLML"

BASE COUNT      1419 a      913 c      1016 g      1413 t
ORIGIN
Query Match      98.5%; Score 4222.8; DB 6; Length 4761;
Best Local Similarity 98.8%; Pred. No. 0;
Matches 4285; Conservative 0; Mismatches 2; Indels 51; Gaps 1;

1  ATGCAAGGCGCAACGCGCTCTGACGCGCGCCCACTCTGCTGTGACGCGCGCGCTG
113  ATGCAAGGCGCAACGCGCTCTGACGCGCGCCCACTCTGCTGTGACGCGCGCGCTG
Qy      61  GCGGTGCTCCCGGCGCTCTGCTGTGACGCGCGCCCACTCTGCTGTGACGCGCGCGCTG
120  GCGGTGCTCCCGGCGCTCTGCTGTGACGCGCGCCCACTCTGCTGTGACGCGCGCGCTG
Db      173  GCGGTGCTCCCGGCGCTCTGCTGTGACGCGCGCCCACTCTGCTGTGACGCGCGCGCTG
232  GCGGTGCTCCCGGCGCTCTGCTGTGACGCGCGCCCACTCTGCTGTGACGCGCGCGCTG
Qy      121  AATGTACTATTGGGGTGAAGCTTCTGGAACAACGCTCTTCAACAGTGAAGTGAAGCG 180
Db      223  AATGTACTATTGGGGTGAAGCTTCTGGAACAACGCTCTTCAACAGTGAAGTGAAGCG 292
Qy      181  GAGCTGCTCAAGACAGCATCAAACTCATCTGCTCTGCTGAGAGCAAGAGAGTCTTT 240
Db      293  GAGCTGCTCAAGACAGCATCAAACTCATCTGCTCTGCTGAGAGCAAGAGAGTCTTT 352
Qy      241  GAAAAAGGCTCTTTAAGACACTTACTCTTCATCACTCACTCTGAGAGTGAAGTGAAG 300
Db      353  GAAAAAGGCTCTTTAAGACACTTACTCTTCATCACTCACTCTGAGAGTGAAGTGAAG 412
Qy      301  ATTATGAGCTACGCTGAACCGGAGCTACCGAGATGAGATTTTATTCTTAATAGTACC 360
Db      413  ATTATGAGCTACGCTGAACCGGAGCTACCGAGATGAGATTTTATTCTTAATAGTACC 472
Qy      361  CGCTTATCATTTGAGACCAAGAAATATCTGCTTCAATCAACAGACAGAGGCTTATAC 420
Db      473  CGCTTATCATTTGAGACCAAGAAATATCTGCTTCAATCAACAGACAGAGGCTTATAC 532
Qy      421  AAGCCAAAGCAAGAGTGAAGTGGATGCTGACCTCTGAGATTTTAAAGCTTAC 480
Db      533  AAGCCAAAGCAAGAGTGAAGTGGATGCTGACCTCTGAGATTTTAAAGCTTAC 592
Qy      481  AAAACCTCTTTAAACATTTCTATTAAAGACCCCAATCAAAATTTGATCAACAGTGTG 540
Db      593  AAAACCTCTTTAAACATTTCTATTAAAGACCCCAATCAAAATTTGATCAACAGTGTG 652
Qy      541  TCACAACAAGTGAATCTTGAAGTCAATTTTCAAACTTTTCAAGCTTATCCCATCAATA 600
Db      653  TCACAACAAGTGAATCTTGAAGTCAATTTTCAAACTTTTCAAGCTTATCCCATCAATA 712
Qy      601  CTGTGTGCTGCTCATTTCAAGTGAATGACGAGATGATATTAATCAATCAATTTAG 660
Db      713  CTGTGTGCTGCTCATTTCAAGTGAATGACGAGATGATATTAATCAATCAATTTAG 772
Qy      661  GTTTCAGATATGATTAACAAAATTTGAAGTGAATTTGACAGACCAATTAATTTGTTCT 720
Db      773  GTTTCAGATATGATTAACAAAATTTGAAGTGAATTTGACAGACCAATTAATTTGTTCT 832

```



QY 721 ATGAATCTAAGCATTTAAATGTACATCAAGGCAAAAGTATACATATGGGAACCCAGTG 780  
 Db 833 ATGAATCTAAGCATTTAAATGTACATCAAGGCAAAAGTATACATATGGGAACCCAGTG 892  
 QY 781 AAAGAGACGTAAGGCTTACATTTTACCTTTATCCTTTTGGGAAAAGAAAATATTT 840  
 Db 893 AAAGAGACGTAAGGCTTACATTTTACCTTTATCCTTTTGGGAAAAGAAAATATTT 952  
 QY 841 ACAAACAATTAAAGATTAATGATCTGCAAACTCTCTTTTAATGATGAAGAGATGAAA 900  
 Db 953 ACAAACAATTAAAGATTAATGATCTGCAAACTCTCTTTTAATGATGAAGAGATGAAA 1012  
 QY 901 AATGTAAATGATCTTCAATATGACCTTCTGCAATACCTGATCTATCTTCCCTGACCA 960  
 Db 1013 AATGTAAATGATCTTCAATATGACCTTCTGCAATACCTGATCTATCTTCCCTGACCA 1072  
 QY 961 GTAGAAATTTTAAACCAAGTACAGATCACTTACAGATATTTCAAGAAATGTAGCACT 1020  
 Db 1073 GTAGAAATTTTAAACCAAGTACAGATCACTTACAGATATTTCAAGAAATGTAGCACT 1132  
 QY 1021 AATGTATCTTCAAGCAATGATTAATCACTGATGTTTTTGAATTAATCACTGCTGTG 1080  
 Db 1133 AATGTATCTTCAAGCAATGATTAATCACTGATGTTTTTGAATTAATCACTGCTGTG 1192  
 QY 1081 AAGCATCTCTCAACTTCAAGCACTGTGAAAGTAACTGCTGCTGATGCGAACCAACTG 1140  
 Db 1193 AAGCATCTCTCAACTTCAAGCACTGTGAAAGTAACTGCTGCTGATGCGAACCAACTG 1252  
 QY 1141 ACTTTGAAGAAAGAAATTAATGTAGTCAATAACAGTACACAGAAACTATATCTGAG 1200  
 Db 1253 ACTTTGAAGAAAGAAATTAATGTAGTCAATAACAGTACACAGAAACTATATCTGAG 1312  
 QY 1201 TACTGAGCGGATCTAACAGTGAAGAAATCAGAAATGGAAGCTTCAAGAAATTAAT 1260  
 Db 1313 TACTGAGCGGATCTAACAGTGAAGAAATCAGAAATGGAAGCTTCAAGAAATTAAT 1372  
 QY 1261 ACTGTCCCCCAAAAGTGAACCTTTTAAGTTGAATTTCCCAATCTCGAGAGATTCAGAGT 1320  
 Db 1373 ACTGTCCCCCAAAAGTGAACCTTTTAAGTTGAATTTCCCAATCTCGAGAGATTCAGAGT 1432  
 QY 1321 CTACAGTTGAAGGCTTATCTTGTGTATGAAGAAAGTACATGATTCATGCTGTTT 1380  
 Db 1433 CTACAGTTGAAGGCTTATCTTGTGTATGAAGAAAGTACATGATTCATGCTGTTT 1492  
 QY 1381 AAGTCTCTAGTGAACATCACTCAATTAACAAGAGATGAAGAAATTAATTAAGTGGGA 1440  
 Db 1493 AAGTCTCTAGTGAACATCACTCAATTAACAAGAGATGAAGAAATTAATTAAGTGGGA 1552  
 QY 1441 TCGCCTTTTGAAGTGTGTGTATGTGCAAAAGATTTGAAGAGATTAAGCTATATGTTA 1500  
 Db 1553 TCGCCTTTTGAAGTGTGTGTATGTGCAAAAGATTTGAAGAGATTAAGCTATATGTTA 1612  
 QY 1501 GTATTCACAGGGAAGATGTGTGTATGTGCAAAAGATTTGAAGAGATTAAGCTATATGTTA 1560  
 Db 1613 GTATTCACAGGGAAGATGTGTGTATGTGCAAAAGATTTGAAGAGATTAAGCTATATGTTA 1672  
 QY 1561 CCAAAAATTTCTTGAATCTCCAAAAGCTGTGTATTTGTATTAATTTGAAGAGATGGG 1620  
 Db 1673 CCAAAAATTTCTTGAATCTCCAAAAGCTGTGTATTTGTATTAATTTGAAGAGATGGG 1732  
 QY 1621 GAAATTTAAAGTGAATGTTCTTAAAAATCTGTGTAGCTGTGTTTAAAAAATTAAGATAAG 1680  
 Db 1733 GAAATTTAAAGTGAATGTTCTTAAAAATCTGTGTAGCTGTGTTTAAAAAATTAAGATAAG 1792  
 QY 1681 CTATATTTGAAGTAAAGTGAAGTGAACCATCTGAGAAAGTCTCTTGAAGATCTGTGT 1740  
 Db 1793 CTATATTTGAAGTAAAGTGAAGTGAACCATCTGAGAAAGTCTCTTGAAGATCTGTGT 1852  
 QY 1741 ACACAGCTGATCTCATATGTTGGATGTGTAGCTGTGCAAAAATGTGGAATCTGATGAAT 1800  
 Db 1853 ACACAGCTGATCTCATATGTTGGATGTGTAGCTGTGCAAAAATGTGGAATCTGATGAAT 1912

QY 1801 GCCTCTAATGATATTAACAATGAAAAATGTGTCCATGATGTGGAACCTTTATTAACAAGCA 1860  
 Db 1913 GCCTCTAATGATATTAACAATGAAAAATGTGTCCATGATGTGGAACCTTTATTAACAAGCA 1972  
 QY 1861 TATTATTTGGGCAAGTGTCAATGATTTCTTTTGAAGTCTTTTCAAGAAATGTGAGCTGGGTA 1920  
 Db 1973 TATTATTTGGGCAAGTGTCAATGATTTCTTTTGAAGTCTTTTCAAGAAATGTGAGCTGGGTA 2032  
 QY 1921 TTGACAGATGCAAACTCAAGAAAGATTAATTAATGATGTGTATTAACAATGAGAAATAT 1980  
 Db 2033 TTGACAGATGCAAACTCAAGAAAGATTAATTAATGATGTGTATTAACAATGAGAAATAT 2092  
 QY 1981 GCTGAAGTGTATTAAGAGAAAAATGAAGCAATATTTGATGATATTTCACTGATCTTTCTTG 2040  
 Db 2093 GCTGAAGTGTATTAAGAGAAAAATGAAGCAATATTTGATGATATTTCACTGATCTTTCTTG 2152  
 QY 2041 GTTAGAGTCCAAATGTCGGAAGAACTTTTCCAGAACTGATGATTTGGCTAGACCAAC 2100  
 Db 2153 GTTAGAGTCCAAATGTCGGAAGAACTTTTCCAGAACTGATGATTTGGCTAGACCAAC 2212  
 QY 2101 ATGGTTTACAGATTTTACCAAGAAATTTGAAGTAACTGTAACCTGATCTATCACTTCTTG 2160  
 Db 2213 ATGGTTTACAGATTTTACCAAGAAATTTGAAGTAACTGTAACCTGATCTATCACTTCTTG 2272  
 QY 2161 GTGGCTACGTGTTTGTGATCTCTGAGAGCTGGGTCTTGACTAAACAATCTCAAGTG 2220  
 Db 2273 GTGGCTACGTGTTTGTGATCTCTGAGAGCTGGGTCTTGACTAAACAATCTCAAGTG 2332  
 QY 2221 GAGCTCAAGCTTCCAAACAATTTTCAATTTTGAATCTTCCCACTGATCTATCAAG 2280  
 Db 2333 GAGCTCAAGCTTCCAAACAATTTTCAATTTTGAATCTTCCCACTGATCTATCAAG 2392  
 QY 2281 GGTAAGAAATTTCTTGAAGAAATTAATTAATTAATTTGAAGATGCACTGAGT 2340  
 Db 2393 GGTAAGAAATTTCTTGAAGAAATTAATTAATTAATTTGAAGATGCACTGAGT 2452  
 QY 2341 AAGTATATCATTAAGAAAGTGAACAAATTTGATATTTCAATGAATCACTCAAGTAAATTAAT 2400  
 Db 2453 AAGTATATCATTAAGAAAGTGAACAAATTTGATATTTCAATGAATCACTCAAGTAAATTAAT 2512  
 QY 2401 GCCACAGGCTCAACAGAGACCTTCTGTGTTCAGAGTGAAGATGGGCACTGTTCTTTT 2460  
 Db 2513 GCCACAGGCTCAACAGAGACCTTCTGTGTTCAGAGTGAAGATGGGCACTGTTCTTTT 2572  
 QY 2461 CCCATCAGGCAACATCTGGGAGAAATTCATATCAAGTCAAGCTCTTTCAACCCT 2520  
 Db 2573 CCCATCAGGCAACATCTGGGAGAAATTCATATCAAGTCAAGCTCTTTCAACCCT 2632  
 QY 2521 GCTTGTGATGTGTCAACCAAGATGATTTTGAAGGCTGAAGAAATGAAGAAATCAAT 2580  
 Db 2633 GCTTGTGATGTGTCAACCAAGATGATTTTGAAGGCTGAAGAAATGAAGAAATCAAT 2692  
 QY 2581 TCAATATCATTTATTAAGTGTGACTGCAATATGCTTACAGATACCTGAAAATCTTTG 2640  
 Db 2693 TCAATATCATTTATTAAGTGTGACTGCAATATGCTTACAGATACCTGAAAATCTTTG 2752  
 QY 2641 AGTTTCAATTTCTCTATATACAGTGTGCAAGTGAAGAAAGTATCAATCACTGCAAT 2700  
 Db 2753 AGTTTCAATTTCTCTATATACAGTGTGCAAGTGAAGAAAGTATCAATCACTGCAAT 2812  
 QY 2701 GGAGATGTTCTTGTGCTTCCATCAATGATGCTTGAAGTATTTGAAGTCTTATGAG 2760  
 Db 2813 GGAGATGTTCTTGTGCTTCCATCAATGATGCTTGAAGTATTTGAAGTCTTATGAG 2872  
 QY 2761 TGTGTGAACAGAAATGATTAATTTTGTCTCCAAATATTTTACATTTTGGATTAATCTGACT 2820  
 Db 2873 TGTGTGAACAGAAATGATTAATTTTGTCTCCAAATATTTTACATTTTGGATTAATCTGACT 2932  
 QY 2821 AAAAAGAAACAACTGACAGATTAATTTGAAGAAAAAGCTCTTCAATTTAGAGCAAGT 2880  
 Db 2933 AAAAAGAAACAACTGACAGATTAATTTGAAGAAAAAGCTCTTCAATTTAGAGCAAGT 2992  
 QY 2881 TACCAAGAGAACTTCTATACAGAGGAAAGATGCTCTTTTCAAGTCTTTTGGAAATTAAT 2940



Db 2993 TACCAAGAGAACTTCTATACAGAGGAGATGCGCTCTTTCAGTCTTTTGGAAATTTAT  
 3052  
 Qy 2941 GACCCCTTGGGAGACCTTGGTGTGAGCTTTTGTAAAGATTTTCCCTGAAGCCGAT  
 3000  
 Db 3053 GACCCCTTGGGAGACCTTGGTGTGAGCTTTTGTAAAGATTTTCCCTGAAGCCGAT  
 3112  
 Qy 3001 CCTTACATAGATATTGATCAGAAATGTGTACACAGAACATACCTTGGCTTAAAGACAT  
 3060  
 Db 3113 CCTTACATAGATATTGATCAGAAATGTGTACACAGAACATACCTTGGCTTAAAGACAT  
 3172  
 Qy 3061 CAGAAATCCAGGCTGAATTTTGGATCCAGAGAGGTGATTCATATGTGAGCTTCAAGGT  
 3120  
 Db 3173 CAGAAATCCAGGCTGAATTTTGGATCCAGAGAGGTGATTCATATGTGAGCTTCAAGGT  
 3232  
 Qy 3121 GGCATTAAGATCAGTAAACCTTACAGGCTTATTTGACTTCTCTCGGGATTTTAA  
 3180  
 Db 3233 GGCATTAAGATCAGTAAACCTTACAGGCTTATTTGACTTCTCTCGGGATTTTAA  
 3292  
 Qy 3181 AAGTATCAGCTTAACTTGAATGTGCAAGAGCTTATTCATTTTGGAGCTGATTCAGT  
 3240  
 Db 3293 AAGTATCAGCTTAACTTGAATGTGCAAGAGCTTATTCATTTTGGAGCTGATTCAGT  
 3352  
 Qy 3241 AAGGAAATTTCAAGCAATTAATCTTACGCTTATTAATCTTATGATTCATCAGTGGG  
 3300  
 Db 3353 AAGGAAATTTCAAGCAATTAATCTTACGCTTATTAATCTTATGATTCATCAGTGGG  
 3412  
 Qy 3301 AGTCCCAAAGCGAAGAGCTTGAATATGCTGACCTTGAAGAGCAACAAAGAGTGGC  
 3360  
 Db 3413 AGTCCCAAAGCGAAGAGCTTGAATATGCTGACCTTGAAGAGCAACAAAGAGTGGC  
 3472  
 Qy 3361 ATGCAATTCGGGTGTCTATCAGAGTCCAAATCTTGTGATCTCCGAGGCAAGCTCCG  
 3420  
 Db 3473 ATGCAATTCGGGTGTCTATCAGAGTCCAAATCTTGTGATCTCCGAGGCAAGCTCCG  
 3532  
 Qy 3421 GAAATTTGAAGTTCAGCTTATGACCTGCTCTCACTTCTTCAATTTGACATTTCTGAG  
 3480  
 Db 3533 GAAATTTGAAGTTCAGCTTATGACCTGCTCTCACTTCTTCAATTTGACATTTCTGAG  
 3592  
 Qy 3481 GGAATTCCTCAATTTGAGGCTGCTGAGAGGCAAGAAAGAAAGTGGTGGTGGTGGT  
 3540  
 Db 3593 GGAATTCCTCAATTTGAGGCTGCTGAGAGGCAAGAAAGAAAGTGGTGGTGGTGGT  
 3652  
 Qy 3541 ACTCAGATACCACTGCTGCTTAAAGGCTGTCTGAAATTTGACAGCCCTTAAATGAATAC  
 3600  
 Db 3653 ACTCAGATACCACTGCTGCTTAAAGGCTGTCTGAAATTTGACAGCCCTTAAATGAATAC  
 3712  
 Qy 3601 GAAAGGACAAATATCCAGTGAACGCTGACGCGGCTTAAAGTCTCAAGTCT-----  
 3651  
 Db 3713 GAAAGGACAAATATCCAGTGAACGCTGACGCGGCTTAAAGTCTCAAGTCT-----  
 3772  
 Qy 3652 -----CTTGGCTGTGTAACGCCA  
 3669  
 Db 3773 CTGATTTGACACACAAACGCTTACTCTTCAACAGACAGAGCTGCTGTGTAACGCCA  
 3832  
 Qy 3670 ATGGCAGTTAATATTTCCGGAATGCTTTGATTTGCTAATTTGCTAGCTCAATGTTGA  
 3729  
 Db 3833 ATGGCAGTTAATATTTCCGGAATGCTTTGATTTGCTAATTTGCTAGCTCAATGTTGA  
 3892  
 Qy 3730 TATATGTGAAGGCTTCTGGGCTTCTTGAAGAGCAGAAATCTTCAATCAAGAGCC  
 3789  
 Db 3893 TATATGTGAAGGCTTCTGGGCTTCTTGAAGAGCAGAAATCTTCAATCAAGAGCC  
 3952  
 Qy 3790 TTTGATTTGATGTTGCTGTTAAAGAAATTAAGATGATCTCAATCATGATTTGAAT  
 3849  
 Db 3953 TTTGATTTGATGTTGCTGTTAAAGAAATTAAGATGATCTCAATCATGATTTGAAT  
 4012  
 Qy 3850 GTGTGTACAGCTTTTCCGGCCCGGGTAGAGTGCAATGCTCTTAATGGAATTAACCTA  
 3909  
 Db 4013 GTGTGTACAGCTTTTCCGGCCCGGGTAGAGTGCAATGCTCTTAATGGAATTAACCTA  
 4072  
 Qy 3910 TTAAGTGCTTTATGCTGCTTCAAGAGCAATTTCTCTGAGGAGACAGTGAAGAAAGT  
 3969

Db 4073 TTAAGTGCTTTATGCTGCTTCAAGAGCAATTTCTTGAAGAGACAGTGAAGAAAGT  
 4132  
 Qy 3970 GAAATGATTCATGAGAAATCTCAACCTTATTTAGATCTGTAATGAAGAACCAAGTTTGT  
 4029  
 Db 4133 GAAATGATTCATGAGAAATCTCAACCTTATTTAGATCTGTAATGAAGAACCAAGTTTGT  
 4192  
 Qy 4030 GTTAAATTTCTGCTGTGAGAAATTTAAAGTTTCAATTAACCAAGATGCTTCAAGTTC  
 4089  
 Db 4133 GTTAAATTTCTGCTGTGAGAAATTTAAAGTTTCAATTAACCAAGATGCTTCAAGTTC  
 4252  
 Qy 4090 ATAGTGATTTACTATGACCAAGAGACAGCGCTGAGAAATTTACTTGAAGTGAAG  
 4149  
 Db 4253 ATAGTGATTTACTATGACCAAGAGACAGCGCTGAGAAATTTACTTGAAGTGAAG  
 4312  
 Qy 4150 CTGCTCTCTGATCCTTGTGAGTATGTCAGAGGCTGCGCTTGTGAGAGATGAGCT  
 4209  
 Db 4313 CTGCTCTCTGATCCTTGTGAGTATGTCAGAGGCTGCGCTTGTGAGAGATGAGCT  
 4372  
 Qy 4210 TCAGGCTCCATCATCATCTTCACTTCACTTTTATTTCTGTTCAAGCTTCTGACTTT  
 4269  
 Db 4373 TCAGGCTCCATCATCATCTTCACTTCACTTTTATTTCTGTTCAAGCTTCTGACTTT  
 4432  
 Qy 4270 ATGGAATTTGGCTGTGA 4287  
 Db 4433 ATGGAATTTGGCTGTGA 4450  
 RESULT 5  
 AK534940  
 LOCUS AK534940 5895 bp DNA linear PAT 22-NOV-2002  
 DEFINITION Sequence 7 from Patent WO02070696.  
 ACCESSION AK534940  
 VERSION AK534940.1 GI:25261486  
 KEYWORDS  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 REFERENCE 1  
 AUTHORS Schuh, A. and Sutherland, R.D.  
 TITLE Cdi09 nucleic acid molecules polypeptides and methods of use  
 JOURNAL Patent: WO 02070696-A 7 12-SEP-2002;  
 Schuh, Andre (CA); Sutherland, Robert D. (CA)  
 FEATURES  
 source location/Qualifiers  
 1..5895  
 /organism="Homo sapiens"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9606"  
 113..4450  
 /note="unnamed protein product"  
 /codon\_start=1  
 /protein\_id="CAD57234.1"  
 /db\_xref="GI:25261487"  
 /translation="MOGPPILTLTAHLCTCTAALAAVAPRFLVTPAGIIRGGNVTL  
 GVLELHCPQVTVKABLLKTAASNLMTVSAEAGVPEKSGFKTLTLPGLPNSADEIV  
 ELRYVGRQDILLNSNTRLSPEKRLSVFQTDKALYKPKQYKFRVTLPSDPKY  
 KTSANLILKDKSNLILQWLSQSDLDGYSKTFQLSHPILGDSIQVQVNDQYQS  
 FQVSRVYLPKEVTLQTPLYCSNMSKHLNGITAKYKGVKVDVTLTFLPLSPWK  
 KKNITKPEKINGSANFSEFNDEKKNVMSGLSEYDLSBPGEVILITVTSYTG  
 SRANSTVYFQHHYITIEFPDYTLVRLKSLNFTAVKTRADGNOLTEBRNNVIT  
 VTORNYTFYMSGNSGNOKMEAVOKINVTYVDSGTFKIEFPILEDSSSLQKAYFLGS  
 KSSNAHSLFSPSKTYIQLKTRBNIKVSGPFLVYSGNRLKELSTMYVSROQVA  
 VGRQSTVSLTPENSTPRKCVIYIYIEDGEIISDVLKIPVQVLFKRLKLTWSKV  
 KAEPSEKSLRISVTPQSVIGIAVADSVNLMANSDITMENVHELELNTGYLIG  
 MENVSFAVQECGLMVLIDANLITQYIDGVNDNAEYERFMEENBGHIVDIDPSLGS  
 SPHAKHPPEFTWIMLDTNMGSRIVQEFVTVDPDITSMVATGPIVISEDLAGLTTTPV  
 ELQAFQPEPIFLNTPYGVIRGEERFALATTIENYKDAITEVVIIEKSPKPIIMTNS  
 IMATGHOOTLLVPSDEGATVLPTRPHLGGIPIYTRALSTPASAATQMLIVABGI  
 EKYSQSTLLDLNRLQSTLKTLSFSPPEVTGSEVQITAGIDVQBSINGASL  
 IRMPYQECBQNMIFAPAIYLLDYLTKKQITLDMKEALSFMOQVQRELLYQREDG  
 SFSAFGVNDPSGTMLSAFVLRCLPADPYIDIDONVLRHTYTWLKGHSGNFGFMD  
 GRVHSELQGNKSPVTLTAYIVSLGKYCOPIVDQESIHFLSEFSRGI SDNYT  
 LALITVALSSVSGPKAKALNMLTWBAQEGCMQFWVSSSEKLSDSMQPRDLDEVAA

Same as  
 Result 1

VALSHFLQFQTSIGIPIMRWLSHQRNSLGGFASOTDTVALKALSEFALMINTERN  
 IOVTVGPSSESPIKFLIDITHNLLLOJAEIAVVPMAVINSAGFEALCOLAVVN  
 VKASGSRRRRSIONOEAPDLVDVAVKENKDINHNVCTSGFSGPERSGMALEVL  
 LSGFVPSSEATLSISETKVEYDHKMLYLDSEVNEOPCNIIPAVNPFVSNTODS  
 VSIQYEPERQAVRSYNSYKLSCOLCSVQCRCEBDGASGSHHSVIFPCKK  
 LHYMEML"

BASE COUNT 1708 a 1174 c 1241 g 1772 t  
 ORIGIN

Query Match 98.5%; Score 4222.8; DB 6; Length 5895;  
 Best Local Similarity 98.8%; Pred. No. 0;  
 Matches 4285; Conservative 0; Mismatches 2; Indels 51; Gaps 1;

Qy 1 ATGAGAGGCCCAACCGCTCTGACCGCGCCCACTCTGCTGCTGCTGACCGCGCGCTG 60  
 Db 113 ATGAGAGGCCCAACCGCTCTGACCGCGCCCACTCTGCTGCTGCTGACCGCGCGCTG 172  
 Qy 61 GCCGTGCTCCCGGAGCTCGGTTTCTGTGACAGCCCCAGGATCATAGGCCCGAGAGA 120  
 Db 173 GCCGTGCTCCCGGAGCTCGGTTTCTGTGACAGCCCCAGGATCATAGGCCCGAGAGA 232  
 Qy 121 AATGACTATTGGGAGGAGCTCTGGAACATGCGCTTCAAGGTGACTGTGAAGGCG 180  
 Db 233 AATGACTATTGGGAGGAGCTCTGGAACATGCGCTTCAAGGTGACTGTGAAGGCG 292  
 Qy 181 GAGCTGCTCAAGACAGATCAAACTCACTGCTCTGCTGAGAGAGAGAGCTCTT 240  
 Db 293 GAGCTGCTCAAGACAGATCAAACTCACTGCTCTGCTGAGAGAGAGAGAGCTCTT 352  
 Qy 241 GAAAAAGGCTCTTTTAAAGACCTTACTCTTCATCACTACTCTGAAAGTGAGATGAG 300  
 Db 353 GAAAAAGGCTCTTTTAAAGACCTTACTCTTCATCACTACTCTGAAAGTGAGATGAG 412  
 Qy 301 ATTATAGGCTACGCTGAACCGGAGCTACCGAGATGAGATTTATCTCTTAATAGTACC 360  
 Db 413 ATTATAGGCTACGCTGAACCGGAGCTACCGAGATGAGATTTATCTCTTAATAGTACC 472  
 Qy 361 CGCTTATCATTTAGACCAAGAGATATCTGCTTCAATTCAGAGCAAGGCTTTATAC 420  
 Db 473 CGCTTATCATTTAGACCAAGAGATATCTGCTTCAATTCAGAGCAAGGCTTTATAC 532  
 Qy 421 AAGCCAAAGCAAGAGTGAAGTTTGCATGTTTACACTCTTCTGAGATTTTAAAGCTTTAC 480  
 Db 533 AAGCCAAAGCAAGAGTGAAGTTTGCATGTTTACACTCTTCTGAGATTTTAAAGCTTTAC 592  
 Qy 481 AAAACCTCTTAAACATTTCTCATTAAGAGCCCCCAATTCAAATTTGATCCAAAGTGTG 540  
 Db 593 AAAACCTCTTAAACATTTCTCATTAAGAGCCCCCAATTCAAATTTGATCCAAAGTGTG 652  
 Qy 541 TCACAAAGAGTGAATCTTGAGTCAATTCAGAACTTTTCACTATCTCCCATCCATA 600  
 Db 653 TCACAAAGAGTGAATCTTGAGTCAATTCAGAACTTTTCACTATCTCCCATCCATA 712  
 Qy 601 CTGTGTGATCTGTCTATTCAGAGTTTCAAGTGAATGACAGACATATATCAATCAATTCAG 660  
 Db 713 CTGTGTGATCTGTCTATTCAGAGTTTCAAGTGAATGACAGACATATATCAATCAATTCAG 772  
 Qy 661 GTTTCAGAAATGATTAACCAAAATTTGAAGTGACTTTTGACAGACACATTAATTTGTTCT 720  
 Db 773 GTTTCAGAAATGATTAACCAAAATTTGAAGTGACTTTTGACAGACACATTAATTTGTTCT 832  
 Qy 721 ATGAATTTTAAAGATTTAAATGTAGCATCAAGCAAGTATATCATATGAGGAGCCAGTG 780  
 Db 833 ATGAATTTTAAAGATTTAAATGTAGCATCAAGCAAGTATATCATATGAGGAGCCAGTG 892  
 Qy 781 AAAGAGACGTAGAGCTTACATTTTATCTTTATCTTTTGGGAAAGAAATATATTT 840  
 Db 893 AAAGAGACGTAGAGCTTACATTTTATCTTTATCTTTTGGGAAAGAAATATATTT 952  
 Qy 841 ACAAACCATTTAAGATTAATGATCTGCAAACTTCTTTTATGATGAAGAGATGAAA 900  
 Db 953 ACAAACCATTTAAGATTAATGATCTGCAAACTTCTTTTATGATGAAGAGATGAAA 1012

Qy 901 AATGTAATGATTCCTTCAATGAGCTTCTGAAATACCTGATCTATCTTCCCTGAGCA 960  
 Db 1013 AATGTAATGATTCCTTCAATGAGCTTCTGAAATACCTGATCTATCTTCCCTGAGCA 1072  
 Qy 961 GTAGAAATTTTAAACACAGTGAACAGATACAGTTACAGTATTTTCAAGAAATGTAACACT 1020  
 Db 1073 GTAGAAATTTTAAACACAGTGAACAGATACAGTTACAGTATTTTCAAGAAATGTAACACT 1132  
 Qy 1021 AATGTGCTCTTCAAGAACATGATTAATCATATGAGTTTGTGATTAATTAATCACTGCTTG 1080  
 Db 1133 AATGTGCTCTTCAAGAACATGATTAATCATATGAGTTTGTGATTAATTAATCACTGCTTG 1192  
 Qy 1081 AAGCCATCTCTCAACTTCAAGCAAGCACTGTGAAGTAACTGTGCTGATGAGCAACAATG 1140  
 Db 1193 AAGCCATCTCTCAACTTCAAGCAAGCACTGTGAAGTAACTGTGCTGATGAGCAACAATG 1252  
 Qy 1141 ACTCTTGAAGAAAGAAATATATGATGCTATTAAGTGAACAGAGAAACCTATCTGAG 1200  
 Db 1253 ACTCTTGAAGAAAGAAATATATGATGCTATTAAGTGAACAGAGAAACCTATCTGAG 1312  
 Qy 1201 TACTGAGCGGATCTAAGAGTGAATCAAGAAATGGAAGCTGTGAGAAATTAATTAAT 1260  
 Db 1313 TACTGAGCGGATCTAAGAGTGAATCAAGAAATGGAAGCTGTGAGAAATTAATTAAT 1372  
 Qy 1261 ACTGCCCCCAAGTGAACCTTTAAGATTTGAATTCCTGAGAGATTCAGTGAAG 1320  
 Db 1373 ACTGCCCCCAAGTGAACCTTTAAGATTTGAATTCCTGAGAGATTCAGTGAAG 1432  
 Qy 1321 CTACAGTTGAAGGCTATTTCTTGTGATGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 1380  
 Db 1433 CTACAGTTGAAGGCTATTTCTTGTGATGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 1492  
 Qy 1381 AAGTCTCTAGTGAACATCATCTCAACTTAAACCAAGAGATGAATTAAGTGAAG 1440  
 Db 1493 AAGTCTCTAGTGAACATCATCTCAACTTAAACCAAGAGATGAATTAAGTGAAG 1552  
 Qy 1441 TCGCCTTTTGAAGTGTGTTTGTAGTGAACCAACCAAGTGAAGTGAAGTGAAGTGAAGT 1500  
 Db 1553 TCGCCTTTTGAAGTGTGTTTGTAGTGAACCAACCAAGTGAAGTGAAGTGAAGTGAAGT 1612  
 Qy 1501 GTATCCAGGGGACAGTTGGTGGCTGTAGAGAAACAAATTTCAACAATGTTCTTTTAA 1560  
 Db 1613 GTATCCAGGGGACAGTTGGTGGCTGTAGAGAAACAAATTTCAACAATGTTCTTTTAA 1672  
 Qy 1561 CGAAGAAATTTCTGAGCTCCAAAGGCTGTGATTTGTGATTAATTTGAAGTGAAGTGG 1620  
 Db 1673 CGAAGAAATTTCTGAGCTCCAAAGGCTGTGATTTGTGATTAATTTGAAGTGAAGTGG 1732  
 Qy 1621 GAATTAATAGTATGTTCTTAAAAATTCCTGTTCACTTGTGTTTAAATTAAGATTAAG 1680  
 Db 1733 GAATTAATAGTATGTTCTTAAAAATTCCTGTTCACTTGTGTTTAAATTAAGATTAAG 1792  
 Qy 1681 CTATATTGAGTAAAGTGAAGTGAACCATCTGAGAAAGTCTCTTTAGAGTCTGTG 1740  
 Db 1793 CTATATTGAGTAAAGTGAAGTGAACCATCTGAGAAAGTCTCTTTAGAGTCTGTG 1852  
 Qy 1741 ACAAGAGCTGACCTCAAGTGTGGAGTTGTAGCTGTGACAAAGTGAATCGATGAAT 1800  
 Db 1853 ACAAGAGCTGACCTCAAGTGTGGAGTTGTAGCTGTGACAAAGTGAATCGATGAAT 1912  
 Qy 1801 GCCTTAATGATTAATCAATGGAATAATGTGTCATGATGGAATCTTTATTAACAGAGA 1860  
 Db 1913 GCCTTAATGATTAATCAATGGAATAATGTGTCATGATGGAATCTTTATTAACAGAGA 1972  
 Qy 1861 TATTAATTAAGCATGTTCAATGAATCTTTTGAAGTCTTTCAAGAAATGTGAGCTCTGGGTA 1920  
 Db 1973 TATTAATTAAGCATGTTCAATGAATCTTTTGAAGTCTTTCAAGAAATGTGAGCTCTGGGTA 2032  
 Qy 1921 TTGACAGATGCAAACTCAAGAGATTAATTTGAAGTGTGTTATGAACAATGAGAAATAT 1980  
 Db 2033 TTGACAGATGCAAACTCAAGAGATTAATTTGAAGTGTGTTATGAACAATGAGAAATAT 2092  
 Qy 1981 GCTGAGAGGTTTATGAGAGAAATGAAGACATATTTGTAGATTAATTCATGACTTTTCTTTG 2040

Db 2093 GCTGAGAGGTTATGAGAGAAATGAGACATATTTGATGATTTATGATACCTTTCTTTG 2152  
Qy 2041 GGTAGAGTCCACATGTCGAAAGATTTTCCAGAGACTTGATTTGGCTAGACACCAAC 2100  
Db 2153 GGTAGAGTCCACATGTCGAAAGATTTTCCAGAGACTTGATTTGGCTAGACACCAAC 2212  
Qy 2101 ATGGGTTACAGATTTTACAGAGATTTGAAAGTAACTGATCTGATCTTCACTCTTGG 2160  
Db 2213 ATGGGTTACAGATTTTACAGAGATTTGAAAGTAACTGATCTGATCTTCACTCTTGG 2272  
Qy 2161 GTGGCTACCTGTTTGTGATCTCTGAGAGCTGGGCTTGGAGCTAACTACTCTCCAGG 2220  
Db 2273 GTGGCTACCTGTTTGTGATCTCTGAGAGCTGGGCTTGGAGCTAACTACTCTCCAGG 2332  
Qy 2221 GAGCTCAAGCCTTCCACCATTTTCACTTTTGAATCTTCCCTACTCTGTTATCAGA 2280  
Db 2333 GAGCTCAAGCCTTCCACCATTTTCACTTTTGAATCTTCCCTACTCTGTTATCAGA 2392  
Qy 2281 GGTGAAGATTTGCTTGGAAATTAATAATTAATTAATTAATTAATTAATTAATTAAT 2340  
Db 2393 GGTGAAGATTTGCTTGGAAATTAATAATTAATTAATTAATTAATTAATTAATTAAT 2452  
Qy 2341 AAGGTATCATTTGAGAAAGTGAACAATTTGATTTCTATGACTTCAAGTGAATTAAT 2400  
Db 2453 AAGGTATCATTTGAGAAAGTGAACAATTTGATTTCTATGACTTCAATGAATTAAT 2512  
Qy 2401 GGCACAGGCGACAGAGAGCCCTTGGTTCCAGAGAGATGGGCGAATGTTCTTTT 2460  
Db 2513 GGCACAGGCGACAGAGAGCCCTTGGTTCCAGAGAGATGGGCGAATGTTCTTTT 2572  
Qy 2461 CCCATAGGCGAACAATCTGGAGAAATTCCTATCAGTCACTCACTCTTCACTCCACT 2520  
Db 2573 CCCATAGGCGAACAATCTGGAGAAATTCCTATCAGTCACTCACTCTTCACTCCACT 2632  
Qy 2521 GCTTCTGATGCTGTCACTCAAGATGATTTTAAAGCTGAGAGAAATGAAGAAATCAT 2580  
Db 2633 GCTTCTGATGCTGTCACTCAAGATGATTTTAAAGCTGAGAGAAATGAAGAAATCAT 2692  
Qy 2581 TCACATTCATCTTATTTAGACTTGACTGACATAGGCTACAGATGACCTCGAAATCTTG 2640  
Db 2693 TCACATTCATCTTATTTAGACTTGACTGACATAGGCTACAGATGACCTCGAAATCTTG 2752  
Qy 2641 AGTTCTCATTTCTCTCTAATACAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 2700  
Db 2753 AGTTCTCATTTCTCTCTAATACAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 2812  
Qy 2701 GAGATGTTCTTGTCTTCCATCAATGCTTACCTCAATGATTTGGAATGCTTATGGC 2760  
Db 2813 GAGATGTTCTTGTCTTCCATCAATGCTTACCTCAATGATTTGGAATGCTTATGGC 2872  
Qy 2761 TGTGTGGAACAGAACTGATTAATTTTGTCCAAATTTTATTAATTTTATGATCT 2820  
Db 2873 TGTGTGGAACAGAACTGATTAATTTTGTCCAAATTTTATTAATTTTATGATCT 2932  
Qy 2821 AAAAAGAAACAACTGACAGATTAATTTGAAGAAAGCTCTTCAATTTATGAGCAAGGT 2880  
Db 2933 AAAAAGAAACAACTGACAGATTAATTTGAAGAAAGCTCTTCAATTTATGAGCAAGGT 2992  
Qy 2881 TACAGAGAGAACTTCTATCAAGAGAGAGATGCTTCTTCACTGCTTTTGGAAATTA 2940  
Db 2993 TACAGAGAGAACTTCTATCAAGAGAGAGATGCTTCTTCACTGCTTTTGGAAATTA 3052  
Qy 2941 GACCTTCTGGAGAGCTTGGTTGTCACTTTTGTATTAAGTGTCTTCTGAGCCGAT 3000  
Db 3053 GACCTTCTGGAGAGCTTGGTTGTCACTTTTGTATTAAGTGTCTTCTGAGCCGAT 3112  
Qy 3001 CCTTACATAGATTTTATGATGAGATGTTTACACAGAACTATCACTGGCTTAAAGCAAT 3060  
Db 3113 CCTTACATAGATTTTATGATGAGATGTTTACACAGAACTATCACTGGCTTAAAGCAAT 3172  
Qy 3061 CAGAAATCCAACTGATTTTGGATTCAGAGAGAGATTTATGATGAGCTTCAAGGT 3120

Db 3173 CAGAAATCCAACTGATTTTGGATTCAGAGAGATGATTCATGATGAGCTTCAAGGT 3232  
Qy 3121 GGCATTAAGATCCAGTAACTTACAGCTTATTTGATTTGATCTCTCTGGGATATAGA 3180  
Db 3233 GGCATTAAGATCCAGTAACTTACAGCTTATTTGATTTGATCTCTCTGGGATATAGA 3292  
Qy 3181 AAGTATCAGCTTACATTTGATGTCAGAGATCTATCCATTTTGGAGTCTGATTCAGT 3240  
Db 3293 AAGTATCAGCTTACATTTGATGTCAGAGATCTATCCATTTTGGAGTCTGATTCAGT 3352  
Qy 3241 AGAGAAATTTGAGAAATTAATCTTACCCCTTAACTTATGATGATTTGATCACTGGGG 3300  
Db 3353 AGAGAAATTTGAGAAATTAATCTTACCCCTTAACTTATGATGATTTGATCACTGGGG 3412  
Qy 3301 AGTCTTAAAGCAAGAGCTTGAATATGCTGACTTGAAGAGAGCAAGAGAGTGGC 3360  
Db 3413 AGTCTTAAAGCAAGAGCTTGAATATGCTGACTTGAAGAGAGCAAGAGAGTGGC 3472  
Qy 3361 ATGCAATTTCTGGGTCTATCAGAGTCCAACTTCTGACTCTGGCAGCCAGCTCCCTG 3420  
Db 3473 ATGCAATTTCTGGGTCTATCAGAGTCCAACTTCTGACTCTGGCAGCCAGCTCCCTG 3532  
Qy 3421 GATATTTGAAGTTCAGCTTATGCTCTCTCACTCTTCACTTCAATTTCACTCTGAG 3480  
Db 3533 GATATTTGAAGTTCAGCTTATGCTCTCTCACTCTTCACTTCAATTTCACTCTGAG 3592  
Qy 3481 GGAATCCCAATTATGAGGTGCTAGAGGCAAGAAATAGCTTGGGTGGTTGATCT 3540  
Db 3593 GGAATCCCAATTATGAGGTGCTAGAGGCAAGAAATAGCTTGGGTGGTTGATCT 3652  
Qy 3541 ACTCAGATATCACTGTGGCTTAAAGGCTGTCTGATTTTGAAGCTTATGATTAATCA 3600  
Db 3653 ACTCAGATATCACTGTGGCTTAAAGGCTGTCTGATTTTGAAGCTTATGATTAATCA 3712  
Qy 3601 GAAAGCAAAATATCCAAAGTGAACGCTGACGGGCTTACTCACTCAAGTCTT----- 3651  
Db 3713 GAAAGCAAAATATCCAAAGTGAACGCTGACGGGCTTACTCACTCAAGTCTT----- 3772  
Qy 3652 -----CTTGGTGGTTCAGGCA 3669  
Db 3773 CTGATTTGACACACACACCGCTTACTCTTCAAGACAGAGCTTGTGTGATAGCA 3832  
Qy 3670 ATGCGATTAATATTTCCGAAATGTTTGGATTTGATTTGATTTGATTTGATTTGAT 3729  
Db 3833 ATGCGATTAATATTTCCGAAATGTTTGGATTTGATTTGATTTGATTTGATTTGAT 3892  
Qy 3730 TATAATGGAAGCTTGTGGGTCTTCTAGAAAGCAAGATCTATCCAAATCAAGAACCC 3789  
Db 3893 TATAATGGAAGCTTGTGGGTCTTCTAGAAAGCAAGATCTATCCAAATCAAGAACCC 3952  
Qy 3750 TTTGATTTAGATTTGCTGTATTAAGAAATTAAGATGATCTCAATCATGTGATTTGAAT 3849  
Db 3953 TTTGATTTAGATTTGCTGTATTAAGAAATTAAGATGATCTCAATCATGTGATTTGAAT 4012  
Qy 3850 GTGTATCAAGCTTTCCGGCCCGGGGTAGAGTGCATGCTTATGGAAGTTATACCTA 3909  
Db 4013 GTGTATCAAGCTTTCCGGCCCGGGGTAGAGTGCATGCTTATGGAAGTTATACCTA 4072  
Qy 3910 TTAAGTGGCTTATGATGCTTCCAGAACCAATTTCTGAGAGAGACAGTGAAGAAATG 3969  
Db 4073 TTAAGTGGCTTATGATGCTTCCAGAACCAATTTCTGAGAGAGACAGTGAAGAAATG 4132  
Qy 3970 GAATATGATCATGGAAGAACTCACTCTATTTAGATTTCTGTAAATGAAGAACCAATTTGT 4029  
Db 4133 GAATATGATCATGGAAGAACTCACTCTATTTAGATTTCTGTAAATGAAGAACCAATTTGT 4192  
Qy 4030 GTTATATTTCTGCTGAGAGAACTTAAAGTTTCAATACCAAGATGCTTCAAGTCTCC 4089  
Db 4133 GTTATATTTCTGCTGAGAGAACTTAAAGTTTCAATACCAAGATGCTTCAAGTCTCC 4252  
Qy 4090 ATAGTGAATTAATATGAGCCAGAGAGACAGGCGGTGAGAAATTAATCACTGAAATGAG 4149  
Db 4253 ATAGTGAATTAATATGAGCCAGAGAGACAGGCGGTGAGAAATTAATCACTGAAATGAG 4312

Query Match	98.5%	Score 4221.4	DB 6	Length 4335
Best local similarity	98.8%	Pred. No. 0		
Matches 4283	Conservative	0	Mismatches	1
			Indels	51
			Gaps	1

QY	1	TTGACAGGGGCCACCGCTCTCGACCGCGGCCACCTCCCTCGCGTGCACCGCGCGCTG	60
Db	1	ATGCAGGGGCCACCGCTCTGCACCGCGGCCACCTCCCTCGCGTGCACCGCGCGCTG	60
QY	61	GCCGTGCTCCCGGGGCTCGGTTTCTGTGTGACAGCCCGAGGATCATCAGCCCGAGGA	120
Db	61	GCCGTGCTCCCGGGGCTCGGTTTCTGTGTGACAGCCCGAGGATCATCAGCCCGAGGA	120
QY	121	AATGTGACTAATTGGGTTGAGCTTCTGGAACATGCGCCCTTACAGGTGACTGTGAAGCG	180
Db	121	AATGTGACTAATTGGGTTGAGCTTCTGGAACATGCGCCCTTACAGGTGACTGTGAAGCG	180
QY	181	GAGCTGCTCAAGACAGCATGAACCTCACTGTCTGTGCTCGGAAGAGAGAGTCTTT	240
Db	181	GAGCTGCTCAAGACAGCATGAACCTCACTGTCTGTGCTCGGAAGAGAGAGTCTTT	240
QY	241	GAAAAAGGCTCTTTTAAAGACACTTACTTTCATCACTACTCTGAA CAGTGCAGATGAG	300
Db	241	GAAAAAGGCTCTTTTAAAGACACTTACTTTCATCACTACTCTGAA CAGTGCAGATGAG	300
QY	301	ATTATATGAGCTACGTGTACCGGACGTACCCAGATGAGATTTTATCTTAATAGTACC	360
Db	301	ATTATATGAGCTACGTGTACCGGACGTACCCAGATGAGATTTTATCTTAATAGTACC	360
QY	361	CGCTTATCATTTGAGACCAAGAAATATCTGTCTTCAATTCAAAACAGCAAGCCCTTATAC	420
Db	361	CGCTTATCATTTGAGACCAAGAAATATCTGTCTTCAATTCAAAACAGCAAGCCCTTATAC	420
QY	421	AAGCCAAAGCAAGAGTGAAGTTTCGATTTGTTCACCTCTTCAGATTTTAAAGCTTAC	480
Db	421	AAGCCAAAGCAAGAGTGAAGTTTCGATTTGTTCACCTCTTCAGATTTTAAAGCTTAC	480
QY	481	AAAACTCTTTTAAACATTTCTCATTTAAGGACCCCAATCAAAATTGATTCACAGTGGTGG	540
Db	481	AAAACTCTTTTAAACATTTCTCATTTAAGGACCCCAATCAAAATTGATTCACAGTGGTGG	540
QY	541	TCACAACAABATGATCTTGGAGTCAATTTCCAAACTTTGACGATCTTCCCATCAATA	600
Db	541	TCACAACAABATGATCTTGGAGTCAATTTCCAAACTTTGACGATCTTCCCATCAATA	600
QY	601	CTTGGTGACTGGTCTAATTCAGATTCAAGTGAATGACAGACATATATTCATCATTTTCAG	660
Db	601	CTTGGTGACTGGTCTAATTCAGATTCAAGTGAATGACAGACATATATTCATCATTTTCAG	660
QY	661	GTTTCAGATATATGATTAACCAAAATTTGAAGTGACTTTGCAGACACCATTAATATGTTCT	720
Db	661	GTTTCAGATATATGATTAACCAAAATTTGAAGTGACTTTGCAGACACCATTAATATGTTCT	720
QY	721	ATGAAATTTAAGCATTTAATATGATACATACCGGCAAAAGTATATATATGGAAGCCAGTG	780
Db	721	ATGAAATTTAAGCATTTAATATGATACATACCGGCAAAAGTATATATATGGAAGCCAGTG	780
QY	781	AAAGGAGACGTAAAGCTTACATTTTACCTTTTGGGGAAGAGAAAAATAT	840
Db	781	AAAGGAGACGTAAAGCTTACATTTTACCTTTTGGGGAAGAGAAAAATAT	840
QY	841	ACAAAAACATTTAAGATTAATGATCTGCAAACTTCTCTTTTAAATGATGAAGATGAAA	900
Db	841	ACAAAAACATTTAAGATTAATGATCTGCAAACTTCTCTTTTAAATGATGAAGATGAAA	900
QY	901	AATGTATATGATTTCTTCAATGACCTTTCGAATACCTGGAATCTATCTTCCCTGAGCA	960
Db	901	AATGTATATGATTTCTTCAATGACCTTTCGAATACCTGGAATCTATCTTCCCTGAGCA	960
QY	961	GTAGAAATTTTAAACACAGTGACAGATCAAGTTACAGTATTTTCAAAATATGAAGACT	1020
Db	961	GTAGAAATTTTAAACACAGTGACAGATCAAGTTACAGTATTTTCAAAATATGAAGACT	1020
QY	1021	AATGTGTCTTCAAGCAACATGATTAATCATATGAGTTTTTGAATTAATCTACTGTCTTG	1080
Db	1021	AATGTGTCTTCAAGCAACATGATTAATCATATGAGTTTTTGAATTAATCTACTGTCTTG	1080
QY	1081	AAGCATCTCTCAACTTCAAGCCACTGTGAGGTAATCTGTGCTGATGCGCAACCAACTG	1140

```
Db 1081 AAGCCATCTCTCAACTTCAAGCCCACTGGAAGTAATCTGTGATGAGCAACCAACG 1140
Qy 1141 ACTCTTGAAGAAAGAAATATATAGTACATTAACAGTACACAGAGAACTATACGAG 1200
Db 1141 ACTCTTGAAGAAAGAAATATATAGTACATTAACAGTACACAGAGAACTATACGAG 1200
Qy 1201 TACTGAGCGGATCTPAACAGTGAATCAAGAAATGGAAGCTGTTCAAGAAATTAAT 1260
Db 1201 TACTGAGCGGATCTPAACAGTGAATCAAGAAATGGAAGCTGTTCAAGAAATTAAT 1260
Qy 1261 ACTGTCCTCCCAAGTGGAACTTTTAAGATTGAATTCCTCAATCTCGAGAGATTCCAGTGA 1320
Db 1261 ACTGTCCTCCCAAGTGGAACTTTTAAGATTGAATTCCTCAATCTCGAGAGATTCCAGTGA 1320
Qy 1321 CTACAGTTGAAGGCTATTTCTGTTAGTAAAGATGAGATGAGGAGTTGATGTCGTT 1380
Db 1321 CTACAGTTGAAGGCTATTTCTGTTAGTAAAGATGAGATGAGGAGTTGATGTCGTT 1380
Qy 1381 AAGTCTCTAGTAAAGATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAAT 1440
Db 1381 AAGTCTCTAGTAAAGATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAAT 1440
Qy 1441 TCGCTTTTGAAGTGTGTTAGTGGCAAAACGATGAGAGATTAACTATATGTA 1500
Db 1441 TCGCTTTTGAAGTGTGTTAGTGGCAAAACGATGAGAGATTAACTATATGTA 1500
Qy 1501 GTATCCAGGGGACAGTGGTGGCTGAGGAAACAAATTCACAAATGTCCTTTAATA 1560
Db 1501 GTATCCAGGGGACAGTGGTGGCTGAGGAAACAAATTCACAAATGTCCTTTAATA 1560
Qy 1561 CCAGAAATTTCTTGAATCTCCAAAGCTGTGTAATGTAATTAATGAAGATGAG 1620
Db 1561 CCAGAAATTTCTTGAATCTCCAAAGCTGTGTAATGTAATTAATGAAGATGAG 1620
Qy 1621 GAAATTAATGATGATGTTTAAATCTCTGTTCACTGTTGTTTAAATTAAGTAAG 1680
Db 1621 GAAATTAATGATGATGTTTAAATCTCTGTTCACTGTTGTTTAAATTAAGTAAG 1680
Qy 1681 CTATATGAGTAAGTAAGTAAGTAAGTAAGTAAGTAAGTAAGTAAGTAAGTAAG 1740
Db 1681 CTATATGAGTAAGTAAGTAAGTAAGTAAGTAAGTAAGTAAGTAAGTAAGTAAG 1740
Qy 1741 ACACAGCCGATCTCAATAGTGGATTTAGCTGTTGCAAAAGTGTGAATCTGAGTA 1800
Db 1741 ACACAGCCGATCTCAATAGTGGATTTAGCTGTTGCAAAAGTGTGAATCTGAGTA 1800
Qy 1801 GCTCTTAATGATATTAACATGGAATGTCATGAGTGGAACTTTAATCAACAGA 1860
Db 1801 GCTCTTAATGATATTAACATGGAATGTCATGAGTGGAACTTTAATCAACAGA 1860
Qy 1861 TATTAATTAAGCATGTTCAATGAATCTTTTGAAGTCTTTCAGGAATGTCAGTGGTA 1920
Db 1861 TATTAATTAAGCATGTTCAATGAATCTTTTGAAGTCTTTCAGGAATGTCAGTGGTA 1920
Qy 1921 TTGACAGATGCAAACTCTCAAGAGATTAATGATGATGTTTATGAACAATGAGATAT 1980
Db 1921 TTGACAGATGCAAACTCTCAAGAGATTAATGATGATGTTTATGAACAATGAGATAT 1980
Qy 1981 GCTGAGAGTTTATGAGAGAAATGAGGACATATGATATTCATGACTTTCTTTG 2040
Db 1981 GCTGAGAGTTTATGAGAGAAATGAGGACATATGATATTCATGACTTTCTTTG 2040
Qy 2041 GGTAGACATCCACATGTCGGAAGCAATTTTCCAGACATTTGGCTTGAACACCAAC 2100
Db 2041 GGTAGACATCCACATGTCGGAAGCAATTTTCCAGACATTTGGCTTGAACACCAAC 2100
Qy 2101 ATGGGTACAGGATTTACCAAGATTTGAAGTAACTGATCTGATCTATCACTCTTGG 2160
Db 2101 ATGGGTACAGGATTTACCAAGATTTGAAGTAACTGATCTGATCTATCACTCTTGG 2160
Qy 2161 GTGGCTACTGTTTGTGATCTCTGAGACCTGGGCTTGAACCACTACCTCAGTG 2220
Db 2161 GTGGCTACTGTTTGTGATCTCTGAGACCTGGGCTTGAACCACTACCTCAGTG 2220

Db 2161 GTGGCTACTGTTTGTGATCTCTGAGACCTGGGCTTGAACCACTACCTCAGTG 2220
Qy 2221 GAGCTCCAGACCTTCCACCAATTTTCAATTTTGAATCTTCCCTACTGTTATCGA 2280
Db 2221 GAGCTCCAGACCTTCCACCAATTTTCAATTTTGAATCTTCCCTACTGTTATCGA 2280
Qy 2281 GGTGAAGATTTGCTTGGAAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2340
Db 2281 GGTGAAGATTTGCTTGGAAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2340
Qy 2341 AAGGTAACTTGAAGAAAGTGAACAAATTTGATATCTTAATGATCTTCAAGTAAATTAAT 2400
Db 2341 AAGGTAACTTGAAGAAAGTGAACAAATTTGATATCTTAATGATCTTCAAGTAAATTAAT 2400
Qy 2401 GCCACAGGCAACACAGACACCTTGTGTCCTGAGTGAAGTGGGCAACTGTTCTTTT 2460
Db 2401 GCCACAGGCAACACAGACACCTTGTGTCCTGAGTGAAGTGGGCAACTGTTCTTTT 2460
Qy 2461 CCATCAAGGCAACACATCTGGAGAAATTCCTATCAAGTCAAGCTCTTCAACCACT 2520
Db 2461 CCATCAAGGCAACACATCTGGAGAAATTCCTATCAAGTCAAGCTCTTCAACCACT 2520
Qy 2521 GCTTCTGATGCTGTCAACCAATGATTTTAAAGGCTGAAGAAATTAATTAATTAAT 2580
Db 2521 GCTTCTGATGCTGTCAACCAATGATTTTAAAGGCTGAAGAAATTAATTAATTAAT 2580
Qy 2581 TCACATCAATCTTATTAATGATGATGATGATGATGATGATGATGATGATGATGAT 2640
Db 2581 TCACATCAATCTTATTAATGATGATGATGATGATGATGATGATGATGATGATGAT 2640
Qy 2641 AGTTCTCATTTCTCTCAATTAACATGATGATGATGATGATGATGATGATGATGAT 2700
Db 2641 AGTTCTCATTTCTCTCAATTAACATGATGATGATGATGATGATGATGATGATGAT 2700
Qy 2701 GAGATGTTCTTGTGCTTCCATCAATGATGATGATGATGATGATGATGATGATGAT 2760
Db 2701 GAGATGTTCTTGTGCTTCCATCAATGATGATGATGATGATGATGATGATGATGAT 2760
Qy 2761 TGTGTGAACAGAAATGATTAATTTGCTCAATTTATTAATTTGATTTGATTTGAT 2820
Db 2761 TGTGTGAACAGAAATGATTAATTTGCTCAATTTATTAATTTGATTTGATTTGAT 2820
Qy 2821 AAAAGAAACAACTGACAGATTAATTTGAAGAAAGCTCTTCAATTAAGGCAAGT 2880
Db 2821 AAAAGAAACAACTGACAGATTAATTTGAAGAAAGCTCTTCAATTAAGGCAAGT 2880
Qy 2881 TACAGAGAGAACTTCTATCAAGAGAGATGATGATGATGATGATGATGATGATGAT 2940
Db 2881 TACAGAGAGAACTTCTATCAAGAGAGATGATGATGATGATGATGATGATGATGAT 2940
Qy 2941 GACCTTCTGGAGACATTTGCTGATGATGATGATGATGATGATGATGATGATGAT 3000
Db 2941 GACCTTCTGGAGACATTTGCTGATGATGATGATGATGATGATGATGATGATGAT 3000
Qy 3001 CTTACATGATTAATGATCAGAAATGTTTGAACAGAAATTAATTAATTAATTAATTAAT 3060
Db 3001 CTTACATGATTAATGATCAGAAATGTTTGAACAGAAATTAATTAATTAATTAATTAAT 3060
Qy 3061 CAGAAATCAACGATGAAATTTTGGATCAGAGAGATGATTAATGATGATGATGAT 3120
Db 3061 CAGAAATCAACGATGAAATTTTGGATCAGAGAGATGATTAATGATGATGATGAT 3120
Qy 3121 GGCATTAAGTCCAGTAACTTAACAGCTTAATTAATTAATTAATTAATTAATTAATTAAT 3180
Db 3121 GGCATTAAGTCCAGTAACTTAACAGCTTAATTAATTAATTAATTAATTAATTAATTAAT 3180
Qy 3181 AAGTATCAAGCTTAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3240
Db 3181 AAGTATCAAGCTTAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3240
Qy 3241 AAGGATCAAGCTTAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3300
Db 3241 AAGGATCAAGCTTAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3300
Qy 3300 AAGGATCAAGCTTAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3360
Db 3300 AAGGATCAAGCTTAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3360
```



OY	3301	AGTCTAAAGGAAAGAAAGCTTTGAAATATGCGACTTGGAGAAGAAACAAGAAAGTGGC	3360
Db	3301	AGCTTAAAGGAAAGAAAGCTTTGAAATATGCGACTTGGAGAAGAAACAAGAAAGTGGC	3360
OY	3361	ATGCATTTCTGGGTGCATCAGAGTCCAACTTCTGACTCTGGCAGCCACGCTCCCTG	3420
Db	3361	ATGCATTTCTGGGTGCATCAGAGTCCAACTTCTGACTCTGGCAGCCACGCTCCCTG	3420
OY	3421	GATATTGAAGTTGCAGCCTATGCACTGCTCTCACACTTCTTACAATTTCAAGACTTCTGAG	3480
Db	3421	GATATTGAAGTTGCAGCCTATGCACTGCTCTCACACTTCTTACAATTTCAAGACTTCTGAG	3480
OY	3481	GGAAATCCCAATTATGAGGGTGGCTAAGCAGGCAAGAAATATGCTGGGGTGGTTTGCATCT	3540
Db	3481	GGAAATCCCAATTATGAGGGTGGCTAAGCAGGCAAGAAATATGCTGGGGTGGTTTGCATCT	3540
OY	3541	ACTCAGAGTATCCACTGCTGGCTTTTAAAGGCTCTGTCGTAATTTGCGCCCTATGAAATACA	3600
Db	3541	ACTCAGAGTATCCACTGCTGGCTTTTAAAGGCTCTGTCGTAATTTGCGCCCTATGAAATACA	3600
OY	3601	GAAGAAGCAAAATATCCAAAGTACCGGTGACGGGGCCTAGCTACCAAGTCTT-----	3651
Db	3601	GAAGAAGCAAAATATCCAAAGTACCGGTGACGGGGCCTAGCTACCAAGTCTTAAAGTTT	3660
OY	3652	-----CTTGGCTGTGTACAGCC	3653
Db	3661	CTGATTTGACACACACAAACCGCTTACTCTCTTACAGACAGACAGATGCTGTGTGTACAGCA	3720
OY	3670	ATGCGAGTTAATATTTCCGCAAAATGGTTTGGATTGCTATTTGTCAGCTCAATGTTGTA	3729
Db	3721	ATGCGAGTTAATATTTCCGCAAAATGGTTTGGATTGCTATTTGTCAGCTCAATGTTGTA	3780
OY	3730	TATAATGTGAAGGCTTCTGGGCTCTTCTAAGAACGAAATCATCCAAATCAAGAACCC	3789
Db	3781	TATAATGTGAAGGCTTCTGGGCTCTTCTAAGAACGAAATCATCCAAATCAAGAACCC	3840
OY	3790	TTTGATTTAGATGTTGCTGTGTAAAGAAATAAAGTATCTCAATCATGTGGAATTTGAT	3849
Db	3841	TTTGATTTAGATGTTGCTGTGTAAAGAAATAAAGTATCTCAATCATGTGGAATTTGAT	3900
OY	3850	GTGCTGTAACAAGTTTGGGGCCCGGGTGTGAGTGCATGGGCTTATGGAAGTTAACTTA	3909
Db	3901	GTGCTGTAACAAGTTTGGGGCCCGGGTGTGAGTGCATGGGCTTATGGAAGTTAACTTA	3960
OY	3910	TTAAGTGCTTATATGTCCTTCAAGAACCAATTTCTGAGCGAGACAGTGAAGAAATG	3969
Db	3961	TTAAGTGCTTATATGTCCTTCAAGAACCAATTTCTGAGCGAGACAGTGAAGAAATG	4020
OY	3970	GAATATGTATCATGAGAAACTCAACCTCTATTATTAGTCTGTAAATGAAGAACCCAGTTTGT	4029
Db	4021	GAATATGTATCATGAGAAACTCAACCTCTATTATTAGTCTGTAAATGAAGAACCCAGTTTGT	4080
OY	4030	GTTAATATTTCTGCTGTGTGAGAAACTTTAAAGTTTCAATATCCCAAGATGCTTCACTGCC	4089
Db	4081	GTTAATATTTCTGCTGTGTGAGAAACTTTAAAGTTTCAATATCCCAAGATGCTTCACTGCC	4140
OY	4090	ATATGAGATTTACTATGAGCCAAAGGAGACAGGGCGGTGAAGATTAACAATCTGAAGTGAAG	4149
Db	4141	ATATGAGATTTACTATGAGCCAAAGGAGACAGGGCGGTGAAGATTTACAATCTGTAAAGTGAAG	4200
OY	4150	CTGTCTCTCTGTGACCTTTGACAGTATGTCCAGGGCTGCCGCTCTTGTGAGAGTGAAGCT	4209
Db	4201	CTGTCTCTCTGTGACCTTTGACAGTATGTCCAGGGCTGCCGCTCTTGTGAGAGTGAAGCT	4260
OY	4210	TCAGGCTCCCATCATCACTCTTCAAGTCAATTTTATTTTCTGTTTCAAGCTTCTGATCTTT	4269
Db	4261	TCAGGCTCCCATCATCACTCTTCAAGTCAATTTTATTTTCTGTTTCAAGCTTCTGATCTTT	4320
OY	4270	ATGGAACCTTGGCTG 4284	
Db	4321	ATGGAACCTTGGCTG 4335	

RESULT 7  
 AX537612 4335 bp DNA linear PAT 23-NOV-2002  
 LOCUS AX537612  
 DEFINITION Sequence 3 from Patent WO02070738.  
 ACCESSION AX537612  
 VERSION AX537612.1 GI:25269535  
 KEYWORDS  
 SOURCE  
 ORGANISM Homo sapiens (human)  
 Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.  
 REFERENCE  
 AUTHORS Schuh, A. and Ouwehand, W.  
 TITLE Diagnosis and treatment of blood disorders  
 JOURNAL Patent: WO 02070738-A 3 12-SEP-2002;  
 Schuh, Andre (CA)  
 FEATURES  
 source Location/Qualifiers  
 1..4335  
 /organism="Homo sapiens"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9606"  
 1..>4335  
 /note="unnamed protein product"  
 /codon\_start=1  
 /protein\_id="CAD57497.1"  
 /db\_xref="GI:25269536"  
 /translation="MGGPPLLTMAHLCTAALAAVAGSPRLVLTARGLIRPGNVITII  
 GVELLHCHPSQVTVKAEILTKTASNLTVSYLAEGVKEKSPKTLITPLSLNAGDEIT  
 ELAVLEHCHPSQVTVKAEILTKTASNLTVSYLAEGVKEKSPKTLITPLSLNAGDEIT  
 KTSILNLIKDPKSNLLOOMLSQSDSLGIVSKTQLSHPLIGMSLOVQNDQTYQS  
 FOYSEVYLPEFEVTLQPTPLCSMSKLNLTITAKYAGKPVAGDVLTLPLSPFKK  
 KKNITKFKINGKNDYFENDEEMKNVMSDNGSEYDLSPGVEILTVTSVTGII  
 SRNVTVTFKPHQDILIEFDYTVTLKPSLNPTATVATRADGNOLTLERRNNVIT  
 VTQRTNRTYMSGNSGQKMAVOKINITYTPQSGTKRIETPILEDSELDKAYELGS  
 KSMNAVHSLKSPSKTYTLQTKDENIKVSPPELVSGNKKLKEKSYVVSQQLVA  
 VGNQNTMSEFLTPENGNTWPACVIVYIIEDDGEIISDVLKI PVQLVEFKNI KLYMSKV  
 KAPSEKSVSLRIIVTQPSDIVGVAAVDSYNLNNASNDIMENYVHELENYGYLGA  
 MPNFSFVFOEGGLAVLTDLANTFKDYIDGVYDAAEYERMEENEGHIVDIDHPSLGS  
 SPVHRKHREPTWIMIDTNMGSRIYOEFEVTPQSIISWANGVAFISDLGLGTLTPV  
 ELQAFQPTPIFLNLPVSVIRGEFALEITIPVILKADTEYVKIIEKDKDILMTSBE  
 INAKHQQTLLVSEDEATVLPFRPHLESLITPVALSPTMSADATWQMLVAREGI  
 EKKYOSILIDLIDNRLQSTLKTLSFSPENVTGSRVQITAIQDVLGSLINLASEI  
 IRRPVGCGEQNMNINFAVNIYLDLTTEKKQITDNLKEKALSFMKQVORIELYORBDQ  
 SFSGFNGLVDSGTSALAFPLRCELEADPYIDIDONVLRHTYMLKQHSNGEFPD  
 GSHVHSLQGNKSPVTLTAVITVLSLIGRKQPNIDVQESIHPLSEFSQSDNYT  
 LALITVLSVSGPKRAEALNMLTWRAEDQGNQPMVSSSKLSDSPORSLDIEVANA  
 IYALVSLPLOTQTSBGLPIKRWLSRQNSLDGFASDTQVTAALKSFLALMNTERTN  
 IQVTVTSGSPVVKFLIDTHNRILLQTAELAVQPAVNIISAGFELACOLNVVN  
 VKSGSSRRRRRSIONCEAPLDVAVKNNKLDLHVNDLVNCTSGSPGRSGALMEVNL  
 LSGMVPSEASLSLETVKEKYEHGKNMLKLDVSNVQFCNIPAVANPKVSNQDASL  
 VSLVDYVPERRAQVRSYNSVEVSKLSDCLSDVQGCRCEDGASGSHHSVIFLCRK  
 LLTFMEMLT"  
 BASE COUNT 1295 a 838 c 922 g 1280 t  
 ORIGIN  
 Query Match 98.4%; Score 4219.8; DB 6; Length 4335;  
 Best Local Similarity 98.8%; Pred. No. 0;  
 Matches 4282; Conservative 0; Mismatches 2; Indels 51; Gaps 1;  
 1 ATGAGAGGAGCCACCGCTCTGCACCGCGCCACCTCTCTGCGTGCACCGCGCGCTG 60  
 1 ATGAGAGGAGCCACCGCTCTGCACCGCGCCACCTCTCTGCGTGCACCGCGCGCTG 60  
 61 GCCCTGCCTCCCGGCGCTCGGTTTCGTGTGACAGCCCGAGGAAATCATCAGCCGGGAGGA 120  
 61 GCCCTGCCTCCCGGCGCTCGGTTTCGTGTGACAGCCCGAGGAAATCATCAGCCGGGAGGA 120  
 121 AATGTGACTATTTGGGGGTGAGAGCTTCTGGAACAACAGCCCTTTCACAGGTGACTGTGAAGCG 180  
 121 AATGTGACTATTTGGGGGTGAGAGCTTCTGGAACAACAGCCCTTTCACAGGTGACTGTGAAGCG 180  
 181 GAGCTGCTCAAGACAGATCAACCTCACTGTCTCTGCTCTGGAAGCAAGAGGACTCTTT 240



Db 181 GAGGCTCAAGACAGATCAAACTCAGCTGCTCTGCTGGAGACAGAGAGCTTT 240  
Qy 241 GAAAAAGGCTCTTTTAAAGACACTTACTCTTCATCACTACCTGTAACAGTGAG 300  
Db 241 GAAAAAGGCTCTTTTAAAGACACTTACTCTTCATCACTACCTGTAACAGTGAG 300  
Qy 301 ATTATGACCTAAGTGTAAACCGACGTAACAGATGAGATTTATCTTAATAGTACC 360  
Db 301 ATTATGACCTAAGTGTAAACCGACGTAACAGATGAGATTTATCTTAATAGTACC 360  
Qy 361 CGCTATATCATTTGAGACCAAGAAATATCTGCTTATTCAAACAGAGGCTTATAC 420  
Db 361 CGCTATATCATTTGAGACCAAGAAATATCTGCTTATTCAAACAGAGGCTTATAC 420  
Qy 421 AAGCCAAAGCAAGTGAAGTTGSCATTTGTAACCTCTCAGATTTTAAAGCTTAC 480  
Db 421 AAGCCAAAGCAAGTGAAGTTGSCATTTGTAACCTCTCTCAGATTTTAAAGCTTAC 480  
Qy 481 AAAACCTCTTTAAACATTTCTCATTAAGACCCCAATCAAAATTTGATCCAGTGTTG 540  
Db 481 AAAACCTCTTTAAACATTTCTCATTAAGACCCCAATCAAAATTTGATCCAGTGTTG 540  
Qy 541 TCACAAACAAAGTATCTTGAAGTCACTTCCAAAATTTTCAAGCTATCTTCCATCCATA 600  
Db 541 TCACAAACAAAGTATCTTGAAGTCACTTCCAAAATTTTCAAGCTATCTTCCATCCATA 600  
Qy 601 CTGAGGACGCTGATTCATCAAGTGAAGTACAGACATATTAATCAATTCAG 660  
Db 601 CTGAGGACGCTGATTCATCAAGTGAAGTACAGACATATTAATCAATTCAG 660  
Qy 661 GTTTCAGATATGATTAACAAATTTGAAGTACTTTCAGACACATTAATTTGTTCT 720  
Db 661 GTTTCAGATATGATTAACAAATTTGAAGTACTTTCAGACACATTAATTTGTTCT 720  
Qy 721 ATGAATTTCAAGATTTTAAATGATACATACAGGCAAGTATCATATGGAAGCCAGTG 780  
Db 721 ATGAATTTCAAGATTTTAAATGATACATACAGGCAAGTATCATATGGAAGCCAGTG 780  
Qy 781 AAAAGGACGTAACGCTTCACTTTTAACTTTTACCTTTTGGGGAAGAAAGAAATAT 840  
Db 781 AAAAGGACGTAACGCTTCACTTTTAACTTTTACCTTTTGGGGAAGAAAGAAATAT 840  
Qy 841 ACAAACATTTTAAATTAATGATCTGCAAACTTCTCTTTTAAATGATGAAGTGA 900  
Db 841 ACAAACATTTTAAATTAATGATCTGCAAACTTCTCTTTTAAATGATGAAGTGA 900  
Qy 901 AATGTAATGATTTCTTCAATGATCTTCTGAAATACCTGATCTATCTTCCCTGAGCA 960  
Db 901 AATGTAATGATTTCTTCAATGATCTTCTGAAATACCTGATCTATCTTCCCTGAGCA 960  
Qy 961 GTAGAAATTTTAAACACAGTGAAGATTAAGTAACTGATTAACAGAAATGTAAGCAT 1020  
Db 961 GTAGAAATTTTAAACACAGTGAAGATTAAGTAACTGATTAACAGAAATGTAAGCAT 1020  
Qy 1021 AATGATCTTCAAGCAATGATTAACATGATGATTTTGAATTAATCTACTGCTTG 1080  
Db 1021 AATGATCTTCAAGCAATGATTAACATGATGATTTTGAATTAATCTACTGCTTG 1080  
Qy 1081 AAGCATCTCTCAACTTCAAGCACTGTAAGTAACTGCTGATGAGCAACCACTG 1140  
Db 1081 AAGCATCTCTCAACTTCAAGCACTGTAAGTAACTGCTGATGAGCAACCACTG 1140  
Qy 1141 ACTCTTGAAGAAAGAAATATGATGATTAACAGTGAACAGAAATCTATCTGAG 1200  
Db 1141 ACTCTTGAAGAAAGAAATATGATGATTAACAGTGAACAGAAATCTATCTGAG 1200  
Qy 1201 TACTGAGGGAATCTTAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 1260  
Db 1201 TACTGAGGGAATCTTAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 1260  
Qy 1261 ACTGTCCTCCAAAGTGAAGCTTTTAAATGATGATTCCAATCTGAGGATTCAGTGA 1320  
Db 1261 ACTGTCCTCCAAAGTGAAGCTTTTAAATGATGATTCCAATCTGAGGATTCAGTGA 1320

Db 1261 ACTGTCCTCCAAAGTGAAGCTTTTAAATGATGATTCCAATCTGAGGATTCAGTGA 1320  
Qy 1321 CTACAGTTGAAGGCTTATTTCTTGTAGTAAAGTGAAGTGAAGTGAAGTGAAGTGA 1380  
Db 1321 CTACAGTTGAAGGCTTATTTCTTGTAGTAAAGTGAAGTGAAGTGAAGTGAAGTGA 1380  
Qy 1381 AAGTCTCTTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 1440  
Db 1381 AAGTCTCTTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 1440  
Qy 1441 TGGCTCTTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 1500  
Db 1441 TGGCTCTTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 1500  
Qy 1501 GTATCAGGAGGACAGTGTGCTGTAAGAAACAAATTTCAACATGCTCTTTAACA 1560  
Db 1501 GTATCAGGAGGACAGTGTGCTGTAAGAAACAAATTTCAACATGCTCTTTAACA 1560  
Qy 1561 CCAGAAATTTCTTGAATCTCAAAAGCTGTGTAATTTGATTAATTAAGATGATGAG 1620  
Db 1561 CCAGAAATTTCTTGAATCTCAAAAGCTGTGTAATTTGATTAATTAAGATGATGAG 1620  
Qy 1621 GAAATTTAATGATGATTTTAAATTTCTGCTTCAAGCTTGTTTTAAATTAAGTGA 1680  
Db 1621 GAAATTTAATGATGATTTTAAATTTCTGCTTCAAGCTTGTTTTAAATTAAGTGA 1680  
Qy 1681 CTATATGGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 1740  
Db 1681 CTATATGGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 1740  
Qy 1741 ACACAGCCTGATCTCAATGTTGGATTTGATGCTTGAACAAAGTGTGAATCTGAAT 1800  
Db 1741 ACACAGCCTGATCTCAATGTTGGATTTGATGCTTGAACAAAGTGTGAATCTGAAT 1800  
Qy 1801 GCTCTTAATGATTAATTAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 1860  
Db 1801 GCTCTTAATGATTAATTAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 1860  
Qy 1861 TATTTATTAAGGATGATTAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 1920  
Db 1861 TATTTATTAAGGATGATTAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 1920  
Qy 1921 TTGACAGATGCAAACTCAAGAGATTAATGATGATTTTGAATTAAGTGAAGTGAAGTGA 1980  
Db 1921 TTGACAGATGCAAACTCAAGAGATTAATGATGATTTTGAATTAAGTGAAGTGAAGTGA 1980  
Qy 1981 GCTGAGGTTTATGAGAGAAATGAAGACATA TTGATGATTAATCACTTTCTTTG 2040  
Db 1981 GCTGAGGTTTATGAGAGAAATGAAGACATA TTGATGATTAATCACTTTCTTTG 2040  
Qy 2041 GGTAGCAGTCCCATGCTCCGAAAGATTTTCCAGAGCTTGAATTTGGCTACACACAA 2100  
Db 2041 GGTAGCAGTCCCATGCTCCGAAAGATTTTCCAGAGCTTGAATTTGGCTACACACAA 2100  
Qy 2101 ATGGGTTACAGATTTTCAAGAAATTTGAAGTAACTGTAACCTGATCTTCTTGG 2160  
Db 2101 ATGGGTTACAGATTTTCAAGAAATTTGAAGTAACTGTAACCTGATCTTCTTGG 2160  
Qy 2161 GTGGCTACTGTTTGTGATCTGAGAGACCTGGGCTTGAATCTTCACTCACTCAAGTG 2220  
Db 2161 GTGGCTACTGTTTGTGATCTGAGAGACCTGGGCTTGAATCTTCACTCACTCAAGTG 2220  
Qy 2221 GAGCTCCAAAGCTTCCAAACATTTTCAATTTTGAATCTTCCCTACTCTGTATCAGA 2280  
Db 2221 GAGCTCCAAAGCTTCCAAACATTTTCAATTTTGAATCTTCCCTACTCTGTATCAGA 2280  
Qy 2281 GGTGAAGATTTGCTTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 2340  
Db 2281 GGTGAAGATTTGCTTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 2340  
Qy 2341 AAGGTAATCATTAAGAAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 2400  
Db 2341 AAGGTAATCATTAAGAAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 2400

QY 2401 GCACAGGCCACACAGACAGCCCTTCTGGTTCACAGTGAAGATGGGGCAACTGTTCTTTT 2460  
 Db 2401 GCACAGGCCACACAGACAGCCCTTCTGGTTCACAGTGAAGATGGGGCAACTGTTCTTTT 2460  
 QY 2461 CCCATCAGGCCACACATCTGGAGAGAAATTCCTATCAAGTCAAGCTCTTCAACCACT 2520  
 Db 2461 CCCATCAGGCCACACATCTGGAGAGAAATTCCTATCAAGTCAAGCTCTTCAACCACT 2520  
 QY 2521 GCTTCTGATGCTGACACAGATGATTTAGTAAAGGCTGAAGAAATGAAATCAAT 2580  
 Db 2521 GCTTCTGATGCTGACACAGATGATTTAGTAAAGGCTGAAGAAATGAAATCAAT 2580  
 QY 2581 TCACATCCTATCTTATAGACTGACATGAGTGAAGTCAAGAGTACCTGAAAATCTTGG 2640  
 Db 2581 TCACATCCTATCTTATAGACTGACATGAGTGAAGTCAAGAGTACCTGAAAATCTTGG 2640  
 QY 2641 AGTTCTCATTTCTCTCTAATACAGTGACTGGCAGTGAAGAGTTCAAGTCACTGCAAT 2700  
 Db 2641 AGTTCTCATTTCTCTCTAATACAGTGACTGGCAGTGAAGAGTTCAAGTCACTGCAAT 2700  
 QY 2701 GGAATGTTCTTGTGCTTCAATGCAATGGCTTACCTCATGATTCGGATCCCTTATGGC 2760  
 Db 2701 GGAATGTTCTTGTGCTTCAATGCAATGGCTTACCTCATGATTCGGATCCCTTATGGC 2760  
 QY 2761 TGTGCTGAACAGAACTATGATTAATTTCTCCAAATATTTACATTTTGAATTCAGT 2820  
 Db 2761 TGTGCTGAACAGAACTATGATTAATTTCTCCAAATATTTACATTTTGAATTCAGT 2820  
 QY 2821 AAAAAAGAAACACAGTACAGATTAATTTGAAGAAAAAGCTCTTCAATTTATGAGCAAGT 2880  
 Db 2821 AAAAAAGAAACACAGTACAGATTAATTTGAAGAAAAAGCTCTTCAATTTATGAGCAAGT 2880  
 QY 2881 TACAGAGAGAACTCTCTATACAGAGGAGATGGCTTTCAAGCTTTTGGGAATAT 2940  
 Db 2881 TACAGAGAGAACTCTCTATACAGAGGAGATGGCTTTCAAGCTTTTGGGAATAT 2940  
 QY 2941 GACCTTCTGGAGACCTTGTGTGACAGTCTTTGTTTGAAGTGTTCCTTGAAGCCGAT 3000  
 Db 2941 GACCTTCTGGAGACCTTGTGTGACAGTCTTTGTTTGAAGTGTTCCTTGAAGCCGAT 3000  
 QY 3001 CCTTACATGATTAATGATCAAGATGTGTTTACACAGAACATACCTTGAAGGACAT 3060  
 Db 3001 CCTTACATGATTAATGATCAAGATGTGTTTACACAGAACATACCTTGAAGGACAT 3060  
 QY 3061 CAGAAATCCACGCTGATTTTGGATCCAGAGAGATGATTCATAGTACCTTCAAGGT 3120  
 Db 3061 CAGAAATCCACGCTGATTTTGGATCCAGAGAGATGATTCATAGTACCTTCAAGGT 3120  
 QY 3121 GGCATTAAGCTCAGTACACTTACAGCTTATATGTAATTTCTCTCTGGATATGAG 3180  
 Db 3121 GGCATTAAGCTCAGTACACTTACAGCTTATATGTAATTTCTCTCTGGATATGAG 3180  
 QY 3181 AAGTATCAGCTTAATGATGTGCAAGAGTCTATCAATTTTGGAGTCTGAATTCAGT 3240  
 Db 3181 AAGTATCAGCTTAATGATGTGCAAGAGTCTATCAATTTTGGAGTCTGAATTCAGT 3240  
 QY 3241 AGAGGAATTTAGACAAATTAATCTTACCTTATATGTAATTTTGGAGTCTGAATTCAGT 3300  
 Db 3241 AGAGGAATTTAGACAAATTAATCTTACCTTATATGTAATTTTGGAGTCTGAATTCAGT 3300  
 QY 3301 AGTCTTAAGGAGGAGGAGCTTTGAAATGCTGAGCTTGAAGGAGGAGCAAGAAAGTGGC 3360  
 Db 3301 AGTCTTAAGGAGGAGGAGCTTTGAAATGCTGAGCTTGAAGGAGGAGCAAGAAAGTGGC 3360  
 QY 3361 ATGCAATTTCTGGGTGTATCAGAGTCCAACTTTCTGACTCTGGACGCAAGCTCTCCCTG 3420  
 Db 3361 ATGCAATTTCTGGGTGTATCAGAGTCCAACTTTCTGACTCTGGACGCAAGCTCTCCCTG 3420  
 QY 3421 GATATTTGAAGTTGACGCTTATGACCTGCTTCAACTTTCAATTTCAAGCTTCTGAG 3480  
 Db 3421 GATATTTGAAGTTGACGCTTATGACCTGCTTCAACTTTCAAGCTTCTGAG 3480

QY 3481 GGAATCCCAATTAATGAGTGTGCTAAGCAGGCAAGAAATAGCTTGGTGTGTTGCACT 3540  
 Db 3481 GGAATCCCAATTAATGAGTGTGCTAAGCAGGCAAGAAATAGCTTGGTGTGTTGCACT 3540  
 QY 3541 ACTCAGATTAACACTGTGGCTTTTAAAGCTCTGTCTAAATTTGAGGCCCTTAATGATCA 3600  
 Db 3541 ACTCAGATTAACACTGTGGCTTTTAAAGCTCTGTCTAAATTTGAGGCCCTTAATGATCA 3600  
 QY 3601 GAAAGCAAAATATCAAGTGAACCGTGAAGGAGGAGCTGATCAACCAATGCT 3651  
 Db 3601 GAAAGCAAAATATCAAGTGAACCGTGAAGGAGGAGCTGATCAACCAATGCT 3651  
 QY 3652 -----CTTGTGTGTATCACCA 3669  
 Db 3652 -----CTTGTGTGTATCACCA 3669  
 QY 3661 CTGATGACACACACACCGCTTACTCTTACAGACAGAGCTTGTGTGTGTACACCA 3720  
 Db 3661 CTGATGACACACACACCGCTTACTCTTACAGACAGAGCTTGTGTGTGTACACCA 3720  
 QY 3720 ATGGCATTAAATATTTCCGCAAAATGTTTGGATTTGCTATTTGCTCAAGTGTGA 3729  
 Db 3720 ATGGCATTAAATATTTCCGCAAAATGTTTGGATTTGCTATTTGCTCAAGTGTGA 3729  
 QY 3721 ATGGCATTAAATATTTCCGCAAAATGTTTGGATTTGCTATTTGCTCAAGTGTGA 3780  
 Db 3721 ATGGCATTAAATATTTCCGCAAAATGTTTGGATTTGCTATTTGCTCAAGTGTGA 3780  
 QY 3730 TATATGTAAGGCTTCTGGGCTTCTAGAGAGCAAGATCTATCCAAATCAAGAGCC 3789  
 Db 3730 TATATGTAAGGCTTCTGGGCTTCTAGAGAGCAAGATCTATCCAAATCAAGAGCC 3789  
 QY 3781 TATATGTAAGGCTTCTGGGCTTCTAGAGAGCAAGATCTATCCAAATCAAGAGCC 3840  
 Db 3781 TATATGTAAGGCTTCTGGGCTTCTAGAGAGCAAGATCTATCCAAATCAAGAGCC 3840  
 QY 3790 TTTGATTTAGATGTTGCTGTAAAGAAATAAAGATGATCTCAATGATGATTTGAAT 3849  
 Db 3790 TTTGATTTAGATGTTGCTGTAAAGAAATAAAGATGATCTCAATGATGATTTGAAT 3849  
 QY 3841 TTTGATTTAGATGTTGCTGTAAAGAAATAAAGATGATCTCAATGATGATTTGAAT 3900  
 Db 3841 TTTGATTTAGATGTTGCTGTAAAGAAATAAAGATGATCTCAATGATGATTTGAAT 3900  
 QY 3900 GTGTGTCAAGCTTTTGGGCGCGGATGAGAGTGGCTCTTATGGAAGTTAACTTA 3909  
 Db 3900 GTGTGTCAAGCTTTTGGGCGCGGATGAGAGTGGCTCTTATGGAAGTTAACTTA 3909  
 QY 3901 GTGTGTCAAGCTTTTGGGCGCGGATGAGAGTGGCTCTTATGGAAGTTAACTTA 3960  
 Db 3901 GTGTGTCAAGCTTTTGGGCGCGGATGAGAGTGGCTCTTATGGAAGTTAACTTA 3960  
 QY 3910 TTAAGTGGCTTTATGAGCTTCTGAGAGCAATTTCTGAGCGAGACAGTGAAGAAAGT 3969  
 Db 3910 TTAAGTGGCTTTATGAGCTTCTGAGAGCAATTTCTGAGCGAGACAGTGAAGAAAGT 3969  
 QY 3961 TTAAGTGGCTTTATGAGCTTCTGAGAGCAATTTCTGAGCGAGACAGTGAAGAAAGT 4020  
 Db 3961 TTAAGTGGCTTTATGAGCTTCTGAGAGCAATTTCTGAGCGAGACAGTGAAGAAAGT 4020  
 QY 3970 GAATATGATCAATGGAATCAACCTTATGATGATCTGTAATGAAACCAAGTTTGT 4029  
 Db 3970 GAATATGATCAATGGAATCAACCTTATGATGATCTGTAATGAAACCAAGTTTGT 4029  
 QY 4021 GAATATGATCAATGGAATCAACCTTATGATGATCTGTAATGAAACCAAGTTTGT 4080  
 Db 4021 GAATATGATCAATGGAATCAACCTTATGATGATCTGTAATGAAACCAAGTTTGT 4080  
 QY 4030 GTTAAATTTCTGTGTGAGAAACTTTAAAGTTTCAATACCAAGTGTCTTCAAGTCTC 4089  
 Db 4030 GTTAAATTTCTGTGTGAGAAACTTTAAAGTTTCAATACCAAGTGTCTTCAAGTCTC 4089  
 QY 4081 GTTAAATTTCTGTGTGAGAAACTTTAAAGTTTCAATACCAAGTGTCTTCAAGTCTC 4140  
 Db 4081 GTTAAATTTCTGTGTGAGAAACTTTAAAGTTTCAATACCAAGTGTCTTCAAGTCTC 4140  
 QY 4090 ATAGTGAATTAATGAGCCAGAGAGCAGCGGTGGAAGTTCAACTCTGAAGTGAAG 4149  
 Db 4090 ATAGTGAATTAATGAGCCAGAGAGCAGCGGTGGAAGTTCAACTCTGAAGTGAAG 4149  
 QY 4141 ATAGTGAATTAATGAGCCAGAGAGCAGCGGTGGAAGTTCAACTCTGAAGTGAAG 4200  
 Db 4141 ATAGTGAATTAATGAGCCAGAGAGCAGCGGTGGAAGTTCAACTCTGAAGTGAAG 4200  
 QY 4150 CTGTCTCTGTGACCTTTGCAAGTATGCAAGGCTGCCGCTTGTGAGATGAGCT 4209  
 Db 4150 CTGTCTCTGTGACCTTTGCAAGTATGCAAGGCTGCCGCTTGTGAGATGAGCT 4209  
 QY 4201 CTGTCTCTGTGACCTTTGCAAGTATGCAAGGCTGCCGCTTGTGAGATGAGCT 4260  
 Db 4201 CTGTCTCTGTGACCTTTGCAAGTATGCAAGGCTGCCGCTTGTGAGATGAGCT 4260  
 QY 4210 TCAGGCTCCATCACTCTTCAAGTATTTTATTTTCTGTTCAAGCTTCTGATCTT 4269  
 Db 4210 TCAGGCTCCATCACTCTTCAAGTATTTTATTTTCTGTTCAAGCTTCTGATCTT 4269  
 QY 4261 TCAGGCTCCATCACTCTTCAAGTATTTTATTTTCTGTTCAAGCTTCTGATCTT 4320  
 Db 4261 TCAGGCTCCATCACTCTTCAAGTATTTTATTTTCTGTTCAAGCTTCTGATCTT 4320  
 QY 4270 ATGGAATTTGGCTG 4284  
 Db 4270 ATGGAATTTGGCTG 4284  
 QY 4321 ATGGAATTTGGCTG 4335  
 Db 4321 ATGGAATTTGGCTG 4335

RESULT 8  
 AX534942  
 LOCUS AX534942 3535 bp DNA linear PAT 22-NOV-2002  
 DEFINITION Sequence 9 from Patent W002070696.  
 ACCESSION AX534942  
 VERSION AX534942.1 GI:25261490  
 KEYWORDS  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1

AUTHORS Schuh, A. and Sutherland, R.D.  
TITLE Cd109 nucleic acid molecules polypeptides and methods of use  
JOURNAL Patent: WO 02070696-A 9 12-SEP-2002;

FEATUES source Location/Qualifiers

1.3535

/organism="Homo sapiens"

/mol\_type="genomic DNA"

/db\_xref="taxon:9606"

113..3316

/note="unnamed protein product"

/codon\_start=1

/protein\_id="CAD57235.1"

/db\_xref="GI:25261491"

/translation="MGGPLLTAAHLICVCTAALAVAPGRLVTAAGIIRPGNVTI  
GVELLEHPSQVTAELKLTASNLTSVLEAGVEKESFFTLTSLPLASADEY  
ELRVTRGODELLFNSRSLSPETKRLSVFIOTDRAKYRQKVRVRYTLPASDEPY  
KTSNLILKIDPKSNLICOVLSQSDLGVSFTFOLSHFILDMSIOVQVNDQTYOS  
FOVSEVLPKFEVTLQTPLYCSMNSHMLGITTAKYTGKPVKGDVTLTFLPSTWKG  
KNITKTEKINGSANFSFDEEMKVMDSNGLSEYLDSPGPEVILTTVESYGI  
SNRVSTNVEFKQHDYIEFFDYTVLKPENLTATKVRADGNOLTEERNNVIT  
VTORNYTEWGSNGSNOKMAVOKINTVPOSGPKIEFPLIDSSLOLKAFLGS  
KSMVAHSLFKSPKTYIOLKTRDENIKVSGPRLVYSGNKKLKEISVWVRGOLVA  
VKQNSMTESLTPENSWTPKACIVYIYIEDDEITSDVLTIVQVFNKIKLTVSKV  
KAESEKSLRISVQPSIVGIIVAVDKSVNLMSANDITMENVAHELNTGYLGS  
MERNFAVEQECGLWLTIDANLTQDYIIDGVNDNAEAEERFMEENGHLYDIDFSLGS  
SPVRKHPEPTEIWLIDNMGYRIYOEFEVLDVSDITSWATGVFSEDLGLITTPV  
ELAOROPFEIPLNTPYSVIRGEPALEITFNVLKIDAREVKYIIEKSDKPELITMNE  
INATGHOCTLVLPSEDDGATVLPPIRPHGEIPITYTALSPPASIDVOTMILVKEI  
EKYSOSILDLIDNRLOSTLTSFSPENVTGSEVQITADIVLSPSINGLASEI  
IRMPYCGGNNMINPANIYILDYLTQKQOLDNLKEXLSPMGQYOEELLYOREDG  
SFSAGNYDPGSGSTWLSAFLVRLCFLEADYIDIDONVLRVYTWLKGHOKSGEEMDP  
GRVISELQGGKSPVTLTAIVTSLDGYRKYQVRFI"

BASE COUNT 1108 a 675 c 731 g 1021 t  
ORIGIN

Query Match 74.4%; Score 3187.4; DB 6; Length 3535;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 3188; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

1 ATGCAAGGCGCCACCGTCTCTGACCGCGCCACCTCTCTGGGTGACCGCGCGCTG 60  
Db 113 ATGCAAGGCGCCACCGTCTCTGACCGCGCCACCTCTCTGGGTGACCGCGCGCTG 172  
Qy 61 GCCGTGGCTCCCGGGGCTCGGTTTCTGCGACAGCCCGAGGATCATCGGCGCGAGGA 120  
Db 173 GCCGTGGCTCCCGGGGCTCGGTTTCTGCGACAGCCCGAGGATCATCGGCGCGAGGA 232  
Qy 121 AATGTACTATTGGGTGAGGCTTCTGAAACACTGCCCTTCAAGGTGACTGTAAGCGG 180  
Db 223 AATGTACTATTGGGTGAGGCTTCTGAAACACTGCCCTTCAAGGTGACTGTAAGCGG 292  
Qy 181 GAGCTGCTCAAGACAGATCAAACTCACTGTCTCTGCTTGGAGCAGAGAGAGCTTTT 240  
Db 293 GAGCTGCTCAAGACAGATCAAACTCACTGTCTCTGCTTGGAGCAGAGAGAGCTTTT 352  
Qy 241 GAAAAAGGCTTTTAAAGACACTTACTCTTCATCACTACTCTGTAAGAGTGAAGTG 300  
Db 353 GAAAAAGGCTTTTAAAGACACTTACTCTTCATCACTACTCTGTAAGAGTGAAGTG 412  
Qy 301 ATTATGAGCTACGTGTAAACCGGACGTACCCAGAGATGAGATTTTATTCTTAATAGTAC 360  
Db 413 ATTATGAGCTACGTGTAAACCGGACGTACCCAGAGATGAGATTTTATTCTTAATAGTAC 472  
Qy 361 CGCTTATCTTTGAAGCAAGAGATATCTGTCTTCAATCAACAGACAGAGCCTTTATAC 420  
Db 473 CGCTTATCTTTGAAGCAAGAGATATCTGTCTTCAATCAACAGACAGAGCCTTTATAC 532  
Qy 421 AAGCCAAAGCAAGAGTGAAGTTGSCATTGTTACCTTCTCAGATTTTTAAAGCCTTAC 480  
Db 533 AAGCCAAAGCAAGAGTGAAGTTGSCATTGTTACCTTCTCAGATTTTTAAAGCCTTAC 592  
Qy 481 AAAACCTTTTAAACATCTCATTAAGAGCCCAATCAAAATTTGATCCAAAGTGTG 540

593 AAAACCTTTTAAACATCTCATTAAGAGCCCAATCAAAATTTGATCCAAAGTGTG 652  
Db 541 TCAGAACAAAGAGATCTTGGAGTCATTTCCAAAACCTTTTACGCTATCTTCCATCCATA 600  
Qy 653 TCAGAACAAAGAGATCTTGGAGTCATTTCCAAAACCTTTTACGCTATCTTCCATCCATA 712  
Db 601 CTGTGTGACTGCTCTATTCAGATTCAGTGAATGACAGACATATTCATCATATTTAG 660  
Qy 713 CTGTGTGACTGCTCTATTCAGATTCAGTGAATGACAGACATATTCATCATATTTAG 772  
Db 661 GTTTTCAAGATATGATATACCAAAATTTGAAGTGAATTCGACACCATTAATTTGTTCT 720  
Qy 773 GTTTTCAAGATATGATATACCAAAATTTGAAGTGAATTCGACACCATTAATTTGTTCT 832  
Db 721 AAGAAATTCAGACATTTAAATGATGCTACACGGAAGTATACATATTCGGAAGCCAG 780  
Qy 833 AAGAAATTCAGACATTTAAATGATGCTACACGGAAGTATACATATTCGGAAGCCAG 892  
Db 781 AAGAGAGAGTAAAGCTTACATTTTACCTTATCTTTGGGGAAGAGAAATATTT 840  
Qy 893 AAGAGAGAGTAAAGCTTACATTTTACCTTATCTTTGGGGAAGAGAAATATTT 952  
Db 841 ACAAAGACATTTAAGATTAATGATCTGCAAACTTCTCTTTTAATGATGAAGATGAA 900  
Qy 953 ACAAAGACATTTAAGATTAATGATCTGCAAACTTCTCTTTTAATGATGAAGATGAA 1012  
Db 901 AATGTAAATGATCTTCAAAATGATCTTCTGAAATCTGATCTATCTTCCCTGAGCA 960  
Qy 1013 AATGTAAATGATCTTCAAAATGATCTTCTGAAATCTGATCTATCTTCCCTGAGCA 1072  
Db 961 GTAGAAATTTTACACAGTGAACAGATCAGGATTAAGGATTAAGAAATGAAGCACT 1020  
Qy 1073 GTAGAAATTTTACACAGTGAACAGATCAGGATTAAGGATTAAGAAATGAAGCACT 1132  
Db 1021 AATGTGTTCTTCAAGCAATGATTAATCATTTGAGTTTGTGATTATATCTGCTG 1080  
Qy 1133 AATGTGTTCTTCAAGCAATGATTAATCATTTGAGTTTGTGATTATATCTGCTG 1192  
Db 1081 AAGCCATCTCTCACTTCAACAGCCACTGTGAAGGTAATCTGCTGATGGAACCAACTG 1140  
Qy 1193 AAGCCATCTCTCACTTCAACAGCCACTGTGAAGGTAATCTGCTGATGGAACCAACTG 1252  
Db 1141 ACTCTGAAGAAAGAAATTAATGATGATTAATGATGATGATGATGATGATGATG 1200  
Qy 1253 ACTCTGAAGAAAGAAATTAATGATGATTAATGATGATGATGATGATGATGATG 1312  
Db 1201 TACTGAGCGGATCTTACAGTGAATATGAAATGGAAGCTGTTCAAGAAATTAATAT 1260  
Qy 1313 TACTGAGCGGATCTTACAGTGAATATGAAATGGAAGCTGTTCAAGAAATTAATAT 1372  
Db 1261 ACTGTCCCCCAAGTGAACCTTTTAAGATGTAATCCAAATCTGAGAGATTCAGTAG 1320  
Qy 1373 ACTGTCCCCCAAGTGAACCTTTTAAGATGTAATCCAAATCTGAGAGATTCAGTAG 1432  
Db 1321 CTACAGTTGAAGGCTTATTTCTTGTGTGTAATAAATACAGAGGCTCATAGTGTGTT 1380  
Qy 1433 CTACAGTTGAAGGCTTATTTCTTGTGTGTAATAAATACAGAGGCTCATAGTGTGTT 1492  
Db 1381 AAGTCTCTAGTAAGACATATCAACCTTAAAAAACAAGATGAAATATAAAGTGGGA 1440  
Qy 1493 AAGTCTCTAGTAAGACATATCAACCTTAAAAAACAAGATGAAATATAAAGTGGGA 1552  
Db 1441 TCGCTTTTGAAGTGTGTGTAAGGCAAAACGATTTGAAGAGTAAAGCTATATGTA 1500  
Qy 1553 TCGCTTTTGAAGTGTGTGTAAGGCAAAACGATTTGAAGAGTAAAGCTATATGTA 1612  
Db 1501 GTATCAGGAGGACATTTGCTGCTGTAAGAAACAAATTCAGCAATGTTCTTTTACA 1560  
Qy 1613 GTATCAGGAGGACATTTGCTGCTGTAAGAAACAAATTCAGCAATGTTCTTTTACA 1672  
Db 1561 CCAGAAATTTCTGACTCCAAAAGCTGTGTAATTTGTAATTAATGAAGATGATGGG 1620  
Qy

Db	1673	CCAGAAAATCTTGAGCATCCAAAAGCCTGTGTAATTGTGATTAATTGAAGATATGAGG	1732
Qy	1621	GAAATTTAAGTGAATGTTCTPAAAAATTCCTGTTCACTGTTTAAAAATTAAGTAAG	1680
Db	1733	GAAATTTAAGTGAATGTTCTPAAAAATTCCTGTTCACTGTTTAAAAATTAAGTAAG	1792
Qy	1681	CTAATTGAGTAAGTAAGGAGCGAACATCTGAGAAAGTCCTTTGAGATCTCTGG	1740
Db	1793	CTAATTGAGTAAGTAAGGAGCGAACATCTGAGAAAGTCCTTTGAGATCTCTGG	1852
Qy	1741	ACAGAGCCTGATCTCATAGATTGGGATTTGATGCTTGAACAAAAGTGTGATCTGAT	1800
Db	1853	ACAAGCCTGATCTCATAGATTGGGATTTGATGCTTGAACAAAAGTGTGATCTGAT	1912
Qy	1801	GCTCTTAATGATTAATTAAGTAAGAAATGTGTCATGATGAGTGAATTTAAACACAGA	1860
Db	1913	GCTCTTAATGATTAATTAAGTAAGAAATGTGTCATGATGAGTGAATTTAAACACAGA	1972
Qy	1861	TATATTATTAAGCATGTTCAATGAATCTTTTGAGCTCTTCAGGAATGTGACCTGGGA	1920
Db	1973	TATATTATTAAGCATGTTCAATGAATCTTTTGAGCTCTTCAGGAATGTGACCTGGGA	2032
Qy	1921	TTGACAGATGAAACCTCAAGAAAGTATTATGATGTTGTTATATACATGACAAATAT	1980
Db	2033	TTGACAGATGAAACCTCAAGAAAGTATTATGATGTTGTTATATACATGACAAATAT	2092
Qy	1981	GCTGAGAGTTTATGAGAGAAATGAAGACATATTGTAGATATTCATGACTTTTCTTGG	2040
Db	2093	GCTGAGAGTTTATGAGAGAAATGAAGACATATTGTAGATATTCATGACTTTTCTTGG	2152
Qy	2041	GCTGAGAGTTTATGAGAGAAATGAAGACATATTGTAGATATTCATGACTTTTCTTGG	2100
Db	2153	GCTGAGAGTTTATGAGAGAAATGAAGACATATTGTAGATATTCATGACTTTTCTTGG	2212
Qy	2101	ATGGGTTACAGAGATTTTACAAAGATTTGAAGTACTGATCTTCACTCTTGG	2160
Db	2213	ATGGGTTACAGAGATTTTACAAAGATTTGAAGTACTGATCTTCACTCTTGG	2272
Qy	2161	GTGGCTACGTGTTTGTGATCTCTGAGAGCTGGGTTCTTGAGCTAACCTACTCCAGTG	2220
Db	2273	GTGGCTACGTGTTTGTGATCTCTGAGAGCTGGGTTCTTGAGCTAACCTACTCCAGTG	2332
Qy	2221	GAGCTCCAGAGCTTTCCAAACATTTTTCATTTTTCATTTTTCATTTTTCATTTTTCAT	2280
Db	2333	GAGCTCCAGAGCTTTCCAAACATTTTTCATTTTTCATTTTTCATTTTTCATTTTTCAT	2392
Qy	2281	GGTGAGAGATTTTGTGGAATTAATCTAATTCAATTAATTGGAAGATCCACTGAGGT	2340
Db	2393	GGTGAGAGATTTTGTGGAATTAATCTAATTCAATTAATTGGAAGATCCACTGAGGT	2452
Qy	2341	AAGGTAATCATTTGAGAAAGTGAACAAATTTGATATTCTAATGACTTCAAGTAAATTAAT	2400
Db	2453	AAGGTAATCATTTGAGAAAGTGAACAAATTTGATATTCTAATGACTTCAAGTAAATTAAT	2512
Qy	2401	GCCTCAGAGCCACGAGAGACCTTCTGTTTCCAGTGAAGATGGGGCAACTGTTCTTTT	2460
Db	2513	GCCTCAGAGCCACGAGAGACCTTCTGTTTCCAGTGAAGATGGGGCAACTGTTCTTTT	2572
Qy	2461	CCCATCAGAGCCACGAGAGACCTTCTGTTTCCAGTGAAGATGGGGCAACTGTTCTTTT	2520
Db	2573	CCCATCAGAGCCACGAGAGACCTTCTGTTTCCAGTGAAGATGGGGCAACTGTTCTTTT	2632
Qy	2521	GCTTCTGATGCTGTCAACCCAGATGATTTTGTAAAGGCTGAAGGATAGAAAATCATAT	2580
Db	2633	GCTTCTGATGCTGTCAACCCAGATGATTTTGTAAAGGCTGAAGGATAGAAAATCATAT	2692
Qy	2581	TCACAAATTCATCTTATTAGACTGACGACATAGGCTACAGAGTACCTGAAAACCTTGG	2640
Db	2693	TCACAAATTCATCTTATTAGACTGACGACATAGGCTACAGAGTACCTGAAAACCTTGG	2752
Qy	2641	AGTTTCTCATTTTCTCTTAATACATGACGACGAGTAAGAAAGTTCAGATCTGCAAT	2700
Db	2753	AGTTTCTCATTTTCTCTTAATACATGACGACGAGTAAGAAAGTTCAGATCTGCAAT	2812

Qy	2701	GGAGATGTTCTTGGTCTTCCATGCAATGGCTTAAGCTCATTTGATTTGGATGCCCTTATGCC	2760
Db	2813	GAGATGTTCTTGGTCTTCCATGCAATGGCTTAAGCTCATTTGATTTGGATGCCCTTATGCC	2872
Qy	2761	TGTGTGAGACAGAACATGATTAATTTTCTCCAAATATTTCATTTTGGATTTACTACT	2820
Db	2873	TGTGTGAGACAGAACATGATTAATTTTCTCCAAATATTTCATTTTGGATTTACTACT	2932
Qy	2821	AAAAAGAAACAATGACATTAATTTGAAAGAAAAAGCTCTTTCATTTATGAGCAAGCT	2880
Db	2933	AAAAAGAAACAATGACATTAATTTGAAAGAAAAAGCTCTTTCATTTATGAGCAAGCT	2992
Qy	2881	TACCAAGAGAACTTCTCATATGAGGGAAGATGGCTCTTTCAGTCTTTGGGAATAT	2940
Db	2993	TACCAAGAGAACTTCTCATATGAGGGAAGATGGCTCTTTCAGTCTTTGGGAATAT	3052
Qy	2941	GACCTTCTGGGAGACCTTGGTTCAGCTTTTGTAAAGATGTTTCTTGAAGCCGAT	3000
Db	3053	GACCTTCTGGGAGACCTTGGTTCAGCTTTTGTAAAGATGTTTCTTGAAGCCGAT	3112
Qy	3001	CCTTACATGATATTGATCAGATGTGTTACACAAACATACACTGGCTTAAAGCAT	3060
Db	3113	CCTTACATGATATTGATCAGATGTGTTACACAAACATACACTGGCTTAAAGCAT	3172
Qy	3061	CAGAAATCCAGGCTGAATTTTGGGATCCAGAGAGATTTCAATGATGACTTCAAGCT	3120
Db	3173	CAGAAATCCAGGCTGAATTTTGGGATCCAGAGAGATTTCAATGATGACTTCAAGCT	3232
Qy	3121	GGCAATTAAGTCCAGTAAACATTTACAGCTTAATTTGTAACTTCTCTCGGATATAGA	3180
Db	3233	GGCAATTAAGTCCAGTAAACATTTACAGCTTAATTTGTAACTTCTCTCGGATATAGA	3292
Qy	3181	AAGTATCAG 3189	
Db	3293	AAGTATCAG 3301	

RESULT 9	AX534944	3535 bp	DNA	linear	PAT 22-NOV-2002
LOCUS	AX534944				
DEFINITION	Sequence 11 from Patent WO02070696.				
ACCESSION	AX534944				
VERSION	AX534944.1	GI:25261494			
KEYWORDS					
SOURCE					
ORGANISM	Homo sapiens (human)				
REFERENCE	1	Schuh, A. and Sutherland, R. D.			
AUTHORS					
TITLE	Cd109 nucleic acid molecules and methods of use				
JOURNAL	Patent: WO 02070696-A 11-SEP-2002.				
FEATUSES					
source					
1..3535					
/organism="Homo sapiens"				</	
/mol_type="genomic DNA"					
/db_xref="taxon:9606"					
113..3316					
/note="unnamed protein product"					
/codon_start=1					
/protein_id="CAD57236.1"					
/db_xref="GI:25261495"					
/translation="MGGPDLTAAHLVCYTRALAVAPGRPLVTARGIIRPGNAVTI					
GYELHEPQSYTYAAELKLTASVLYSVLEAGVPEKSKLTLPISLPLNSADELY					
ELKVTGRQDEILVSNSTRLSFETKRISVFIQIDKALKPKOEKRIIVLPSDFKY					
KLSNLILIKDPSKSLIQWLSQSDLEVTISFTPLSSHPIUGDWSIYOVDNDITYOS					
FOVSEYVLKFEKAVTLQTPYCSMNSKHLNGITIAKTYGKPVKQDVTLPISLPSFWGK					
KNILKTEPKVSNANFSFNDEKMNKWSNGSEYLDLSPGVEILITVTESVYGI					
SRNVSTNVPFKOHVYIEFPDYTVLAKPSNFTATVYTAGDGNQLTEBRNNVYT					
VYQRVYTYKMSGNGKMAVOKINYTPVDSCTFKIEPPIILDISSLEOLKAYPGS					
KSSMAVHSPLFSPSKTYIQLKTRDENTIKVSPFLVVSQNKRLKELSYNVVSKQLYVA					

OY		2701	GGAGATGTTCTTGGTCCTTGCCATCAAAAGCGTTAGCCTCATTTGATGTGGAGTGCCTTAATGCC	2760
Db		2813	GGAGATGTTCTTGGTCCTTGCCATCAAAAGCGCTTAGCCCTCATTTGATGTGGAGTGCCTTAATGCC	2872
OY		2761	TGTGGTGAACAGAACAATGATAAATTTTGTGCTCAAATATTATACATTTTGGATTACTGACT	2820
Db		2873	TGTGGTGAACAGAACAATGATAAATTTTGTGCTCAAATATTATACATTTTGGATTACTGACT	2932
OY		2821	AAAAAGAAACAACATGACAGATAATTGAAAAGAAAAGAGCTCTTTCATTATATGAGGCAAGGT	2880
Db		2933	AAAAAGAAACAACATGACAGATAATTGAAAAGAAAAGAGCTCTTTCATTATATGAGGCAAGGT	2992
OY		2881	TACCAAGAGAACCTTCTCTATTCAGAGGGAAGATGCGCTTTCAGTGGTTTTGGAAATAT	2940
Db		2993	TACCAAGAGAACCTTCTCTATTCAGAGGGAAGATGCGCTTTCAGTGGTTTTGGAAATAT	3052
OY		2941	GACCCTTCTGGGAGACATTTGGTGTGACGCTTTTGTTTAAGATGTTTCCTTTGAAGCCGAT	3000
Db		3053	GACCCTTCTGGGAGACATTTGGTGTGACGCTTTTGTTTAAGATGTTTCCTTTGAAGCCGAT	3112
OY		3001	CCTTACATAGATATTTGATCAGAAATGTGTTACAGAAACATACACTTGGCTTAAAGACAT	3060
Db		3113	CCTTACATAGATATTTGATCAGAAATGTGTTACAGAAACATACACTTGGCTTAAAGACAT	3172
OY		3061	CAGAAATCCAAACGGTGAAATTTTGGGATCCAGGAAGATGATTCAATAGTAGCTTCAAGGT	3120
Db		3173	CAGAAATCCAAACGGTGAAATTTTGGGATCCAGGAAGATGATTCAATAGTAGCTTCAAGGT	3232
OY		3121	GGCAATAAAGTCAGATACACTTACAGCCCTATATTTGTAACCTCTCTCTGGGATATAGA	3180
Db		3233	GGCAATAAAGTCAGATACACTTACAGCCCTATATTTGTAACCTCTCTCTGGGATATAGA	3292
OY		3181	AAGTATCAG 3189 	
Db		3293	AAGTATCAG 3301 	
RESULT 9				
AX534944		AX534944	3535 bp DNA linear	PAT 22-NOV-2002
LOCUS		Sequence 11 from Patent WO02070696.		
DEFINITION		AX534944		
ACCESSION		AX534944.1 GI:25261494		
VERSION				
KEYWORDS		Homo sapiens (human)		
SOURCE		Homo sapiens		
ORGANISM		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
REFERENCE		1 Schuh, A. and Sutherland, R.D. Cd109 nucleic acid molecules polypeptides and methods of use Patent: WO 02070696-A 11 12-SEP-2002;		
AUTHORS		Schuh, Andre (CN) ; Sutherland, Robert D. (CA)		
TITLE		Location/Qualifiers		
JOURNAL		1. .3535 /organism="Homo sapiens" /mol_type="genomic DNA" /db_xref="taxon:9606"		
FEATURES		113..3316 /note="unnamed protein product"		
SOURCE		/codon_start=1 /protein_id="CAD57236.1" /db_xref="GI:25261495"		
CDS		/translation="MOGPPLLTAHAHLVCVTALAAVAPGRPLVTAPGIRPGANVTII GVLELHCPSQIVKAEILKTASNLTVSVTEAGVCFEKSFRKLTLPIPLNSADELY ELRVTGRTDEILLFNSSTRLSFEETKAISVTFOTDKALKYKOEVKRAITVLPSDFRY KTSINTLIKDPKSNLIQQWLSQSDLGAVISKTPQLSSHPLIGWSIIOVONDVDTYOS FOVESEVYLPRFEVTLQTPLYSCKMSKHNLNTITAKTYGSPVGAVTLTPLPSFMWK KNKITKFKINGANSFENDENKMSNDNGSEVYLDLSSPPEVILTFTVESLVGI SRNVSTNVPFKODYIEEFDDYTTLKPSLNPAFYATKADNONLTLEERRNNVVIT VTORATTEVWGSNSGNQMEVAOKINKITVLPDGTFRKEIPRLIEDSSELQKAYFGVS KSSNAVHSLEKSPSKTIYIQTKTDENIKGPSFELVSGNKRLKEISYNVSVSQGLVA		

VGKQNSTHPSLTPENSWTPKACVIVYIEDGEIISDVLIKIPVQVFNKIKLWYSKV  
 KAEPSKXSLRISTVOPDSIVGIIVANDKSVNLMASNDITMENVENHIELVNTGYLGS  
 MEMNSFAVECEGLWVLDJANLTKDYIGCVNDAVYAEEMENVEHIELVNDIDFSLGS  
 SPVAKHPPEETIMLMDJNMGSRILYOEFEVTPDSITSWATGFIYSBDLGGGLTTTPV  
 ELQAROPPEIPLNLPYSVRGESEFALEITIENYLDARBEVRIIEKSDKFDLMTSNE  
 INATHQOOLVPSDEGATVLEPIRPTHIGEIPITVTALSPASDAVOMLILVKEGI  
 EKSYISQILIDJNRLQSLTKLSFSPPTNTVGSERQJITAIIDVLPISINGLASL  
 IRMPYGCGBOMINFAPNYIIDLVLTKKKOLDNLKEXALFMRGQYRELLYOREDS  
 SFSAFNGDPSGSTWLSAFVLCFLEADYIDIDONVLFHYTWMLKQHSNGEFWMD  
 GRVISEHLOGKNSPVTLLAYIVTSLGKRYQVFRIR

BASE COUNT 1107 a 676 c 731 g 1021 t  
 ORIGIN

Query Match 74.3%; Score 3185.8; DB 6; Length 3535;  
 Best Local Similarity 99.9%; Pred. No. 0;  
 Matches 3187; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

1 ATGCAGGGCCACCGCTCTGACCGCGCCACCTCTCTGCGTGTGACCGCGCGCTG 60  
 113 ATGCAGGGCCACCGCTCTGACCGCGCCACCTCTCTGCGTGTGACCGCGCGCTG 172  
 61 GCGGTGGCTCCCGGGGCTGGGTTCTGCTGACAGCCCGGAGATCATCGGCCGAGA 120  
 173 GCGGTGGCTCCCGGGGCTGGGTTCTGCTGACAGCCCGGAGATCATCGGCCGAGA 232  
 121 AATGTACTATTGGGGTGGAGCTTCTGAAACACTGCGCTTCAAGGTGACTGAAAGCG 180  
 233 AATGTACTATTGGGGTGGAGCTTCTGAAACACTGCGCTTCAAGGTGACTGAAAGCG 292  
 181 GAGCTGCTCAAGACAGCATCAAACTCACTGCTCTGCTCTGAGAACAGAGAGCTTT 240  
 293 GAGCTGCTCAAGACAGCATCAAACTCACTGCTCTGCTCTGAGAACAGAGAGCTTT 352  
 241 GAAAAAGGCTCTTTAAGACACTTACTCTTCATCACTACCTCTGAAAGTGCAGATGAG 300  
 353 GAAAAAGGCTCTTTAAGACACTTACTCTTCATCACTACCTCTGAAAGTGCAGATGAG 412  
 301 ATTATAGACTAGCTGCTGTAACCGGACGTAACCGAGATGAGATTTTATTCTTAATGTAAC 360  
 413 ATTATAGACTAGCTGCTGTAACCGGACGTAACCGAGATGAGATTTTATTCTTAATGTAAC 472  
 361 GCGTTATCTTTGAGACCAAGAAATATCTGCTTCATTTCAACAGACAGAGCTTATAC 420  
 473 GCGTTATCTTTGAGACCAAGAAATATCTGCTTCATTTCAACAGACAGAGCTTATAC 532  
 421 AAGCCAAAGCAAGAGTGAAGTTGGCATTTGTTACACTTTCTCAGATTTTAAGCTTAC 480  
 533 AAGCCAAAGCAAGAGTGAAGTTGGCATTTGTTACACTTTCTCAGATTTTAAGCTTAC 592  
 481 AAAAACTCTTTAAACATTTCTCAATTAAAGACCCCAATCAATTTGATCCAAAGGTG 540  
 593 AAAAACTCTTTAAACATTTCTCAATTAAAGACCCCAATCAATTTGATCCAAAGGTG 652  
 541 TCACACAAAGTGAATCTTGAGTCAATTTCCAAAATTTTCACTATCTTCCATCCATA 600  
 653 TCACACAAAGTGAATCTTGAGTCAATTTCCAAAATTTTCACTATCTTCCATCCATA 712  
 601 CTTCGTGACTGCTTATTTCAAGTTCAAGTGAATGACACAGACATTTATCATCATTTGAG 660  
 713 CTTCGTGACTGCTTATTTCAAGTGAATGACACAGACATTTATCATCATTTGAG 772  
 661 GTTTCGAAATATGATTTTACCAAAATTTGAAAGTGAATTTGACAGACCATTATATGTTCT 720  
 773 GTTTCGAAATATGATTTTACCAAAATTTGAAAGTGAATTTGACAGACCATTATATGTTCT 832  
 721 ATGAATTTTACGATTTTAAATGATACATCACGGCAAGATATACATATGGAAGCCAGTG 780  
 833 ATGAATTTTACGATTTTAAATGATACATCACGGCAAGATATACATATGGAAGCCAGTG 892  
 781 AAGAGAGAGTAAAGCTTACATTTTACCTTTATCTTTTGGGAAAGAAAGAAATATTT 840  
 893 AAGAGAGAGTAAAGCTTACATTTTACCTTTATCTTTTGGGAAAGAAAGAAATATTT 952

841 AAAAAACATTTAAGATTAATGATCTGCAAACTTCTCTTTAAATGATGAAGATGAAA 900  
 953 AAAAAACATTTAAGATTAATGATCTGCAAACTTCTCTTTAAATGATGAAGATGAAA 1012  
 901 AATGTAATGATTTCTTCAATGAGACTTTCTGAATTCCTGATCTATCTTCCCTGAGACA 960  
 1013 AATGTAATGATTTCTTCAATGAGACTTTCTGAATTCCTGATCTATCTTCCCTGAGACA 1072  
 961 GTAGAAATTTTAACACAGTGAACGAATCAGTTACAGTATTTTCAAGAAATGAGACT 1020  
 1073 GTAGAAATTTTAACACAGTGAACGAATCAGTTACAGTATTTTCAAGAAATGAGACT 1132  
 1021 AATGTGTTCTTCAAGCAACATGATTAACATGAGTTTGTGATTTTGAATTAATACTGCTTG 1080  
 1133 AATGTGTTCTTCAAGCAACATGATTAACATGAGTTTGTGATTTTGAATTAATACTGCTTG 1192  
 1081 AAGCATCTCTCAACTTCAAGCAACCTGTAAGGTTACTGCTGATGGAACCAACTG 1140  
 1193 AAGCATCTCTCAACTTCAAGCAACCTGTAAGGTTACTGCTGATGGAACCAACTG 1252  
 1141 ACTCTGAAGAAAGAAAGAAATATGATGCTATTAACAGTGAACAGAGAAACTATATGAG 1200  
 1253 ACTCTGAAGAAAGAAAGAAATATGATGCTATTAACAGTGAACAGAGAAACTATATGAG 1312  
 1201 TACTGAGCGGATCTTAACAGTGAAGATGAGAAATGAGAACTGTTCAAGAAATTAATAT 1260  
 1313 TACTGAGCGGATCTTAACAGTGAAGATGAGAAATGAGAACTGTTCAAGAAATTAATAT 1372  
 1261 ACTGTCCCCCAAGTGAACCTTTTAAGTGAATTTCCCAATCTGAGAGATTCAGTGAAG 1320  
 1373 ACTGTCCCCCAAGTGAACCTTTTAAGTGAATTTCCCAATCTGAGAGATTCAGTGAAG 1432  
 1321 CTACAGTTGAAGGCTTATTTCTGTTGATGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1380  
 1433 CTACAGTTGAAGGCTTATTTCTGTTGATGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1492  
 1381 AAGTCTCTTGAAGACATACATCACTCAATAAACAAGAGTGAAGTGAAGTGAAGTGAAG 1440  
 1493 AAGTCTCTTGAAGACATACATCACTCAATAAACAAGAGTGAAGTGAAGTGAAGTGAAG 1552  
 1441 TCGCTTTTGAAGTGGTGGTGAAGCAAAAGATGAAGAGTGAAGTGAAGTGAAGTGAAG 1500  
 1553 TCGCTTTTGAAGTGGTGGTGAAGCAAAAGATGAAGAGTGAAGTGAAGTGAAGTGAAG 1612  
 1501 GTATTCAGGGGACAGTGGTGGTGAAGCAAAAGATGAAGAGTGAAGTGAAGTGAAGTGAAG 1560  
 1613 GTATTCAGGGGACAGTGGTGGTGAAGCAAAAGATGAAGAGTGAAGTGAAGTGAAGTGAAG 1672  
 1561 CGAGAAATTTCTGAGACTCAAAAGCTGTGTAATTTGATTAATTTGAAGATGATGG 1620  
 1673 CGAGAAATTTCTGAGACTCAAAAGCTGTGTAATTTGATTAATTTGAAGATGATGG 1732  
 1621 GAAATTTATAGTGAATGTTCTTAAATTTCTGTTCAAGTGTGTTTAAATTAAGTGAAG 1680  
 1733 GAAATTTATAGTGAATGTTCTTAAATTTCTGTTCAAGTGTGTTTAAATTAAGTGAAG 1792  
 1681 CTATTTGAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1740  
 1793 CTATTTGAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1852  
 1741 ACACAGCTGACTCATATGTTGGGATTTGAGTGTGACATGTAAGAGTGAAGTGAAGTGAAGT 1800  
 1853 ACACAGCTGACTCATATGTTGGGATTTGAGTGTGACATGTAAGAGTGAAGTGAAGTGAAGT 1912  
 1801 GCTCTTAATGATTTTACATGAGAAATGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1860  
 1913 GCTCTTAATGATTTTACATGAGAAATGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1972  
 1861 TATTTATTTAGGAGTTCATGAATTTCTTTGACGCTTTTCAAGAAATGAGTGAAGTGAAGTGAAG 1920  
 1973 TATTTATTTAGGAGTTCATGAATTTCTTTGACGCTTTTCAAGAAATGAGTGAAGTGAAGTGAAG 2032  
 1921 TTGACAGATGCAAACTTCAGAGAGATTAATATGATGAGTGTGTTTGAACAAATGAGAAATAT 1980



D	2033	TTGACAGATGCAAACTTCACAGAGATTAATGATGCTTTTATGCAAGCAGAAATAT	2092	D	3113	CTTACATAGATATTGATCAGAAATGCTTACACAGAAACATACCTGGCTAAAGACAT	3172
Q	1981	GCTAGAGGTTTATGAGAGAAATGAGAGACATATTGATATTATTCAGACTTTCTTTG	2040	Q	3061	CAGAAATCCAAACGGTGAATTTGGGATCCAGAAAGTGAATTGATAGTCAAGT	3120
D	2093	GCTAGAGGTTTATGAGAGAAATGAGAGACATATTGATATTATTCAGACTTTCTTTG	2152	D	3173	CAGAAATCCAAACGGTGAATTTGGGATCCAGAAAGTGAATTGATAGTCAAGT	3232
Q	2041	GATAGCACTCCACATGTCGGAAGCAATTTTCAGAGACTTGGATTGGCTAGACCAAC	2100	Q	3121	GGCAATPMAAGTCCAGTACACTTACAGCCTATATTGTAACCTTCTCTGGGATYAGA	3180
D	2153	GATAGCACTCCACATGTCGGAAGCAATTTTCAGAGACTTGGATTGGCTAGACCAAC	2212	D	3233	GGCAATPMAAGTCCAGTACACTTACAGCCTATATTGTAACCTTCTCTGGGATYAGA	3292
Q	2101	ATGGCTTACAGATTTTACCAAGATTGTAAGTAACTGACTGATCTATCACTTTCTGG	2160	Q	3181	AAGTATCAG 3189	
D	2213	ATGGCTTACAGATTTTACCAAGATTGTAAGTAACTGACTGATCTATCACTTTCTGG	2272	D	3293	AAGTATCAG 3301	
Q	2161	GTGGCTACTGTTTGTGATCTGAGAGACCTGGGCTTGGAGTAAACAACCTACTCCAGT	2220	RESULT 10			
D	2273	GTGGCTACTGTTTGTGATCTGAGAGACCTGGGCTTGGAGTAAACAACCTACTCCAGT	2332	AY083458 4735 bp mRNA linear ROD 08-Apr-2002			
Q	2221	GAGCTCCAAAGCCTTCCAAACATTTTTCATTTTGTGATCTTCCCTACTCTGTTATCAGA	2280	LOCUS AY083458			
D	2233	GAGCTCCAAAGCCTTCCAAACATTTTTCATTTTGTGATCTTCCCTACTCTGTTATCAGA	2392	DEFINITION Mus musculus GPI-anchored alpha-2 macroglobulin-related protein			
Q	2281	GCTGAGAAATTTGCTTTGGAATACTATATTCAATTTTGAAGAATGCCATGAGGT	2340	ACCESSION AY083458			
D	2393	GCTGAGAAATTTGCTTTGGAATACTATATTCAATTTTGAAGAATGCCATGAGGT	2452	VERSION AY083458.1 GI:20070079			
Q	2341	AAGGTAATCATTTGAAAGAGCAAAATTTGATTTTCATAGACTTCAAGTAATAAT	2400	KEYWORDS			
D	2453	AAGGTAATCATTTGAAAGAGCAAAATTTGATTTTCATAGACTTCAAGTAATAAT	2512	SOURCE			
Q	2401	GCCAAGGCCACAGACAGACCTTCTGGTCCAGTGGAGTGGGCGCACTGTTCTTTT	2460	ORGANISM			
D	2513	GCCAAGGCCACAGACAGACCTTCTGGTCCAGTGGAGTGGGCGCACTGTTCTTTT	2512	REFERENCE			
Q	2461	CCCATCAGGCCCAACACATCTGGGAGAAATTCCTATCAGAGTCAACGCTCTTCAACCACT	2520	JOURNAL			
D	2573	CCCATCAGGCCCAACACATCTGGGAGAAATTCCTATCAGAGTCAACGCTCTTCAACCACT	2632	REFERENCE			
Q	2521	GCTTCTGATGCTGTCACCCAGATGATTTAGTAAAGGTGAAGAAATGAAATCATAT	2580	AUTHORS			
D	2633	GCTTCTGATGCTGTCACCCAGATGATTTAGTAAAGGTGAAGAAATGAAATCATAT	2632	TITLE			
Q	2581	TCACATATCCATCTTATTAGACTGACATAGAGCTACAGATACCCCTGAAACTTTG	2640	JOURNAL			
D	2693	TCACATATCCATCTTATTAGACTGACATAGAGCTACAGATACCCCTGAAACTTTG	2752	REFERENCE			
Q	2641	AGTTTCTCATTTCTCTTAATACAGTGACTGGCAGTGAAGAGTTCAAGATCACTCAATT	2700	AUTHORS			
D	2753	AGTTTCTCATTTCTCTTAATACAGTGACTGGCAGTGAAGAGTTCAAGATCACTCAATT	2812	TITLE			
Q	2701	GGAGATGTTCTTGGTCCCTTCATCAATGCTTACCTCATTTGCGAGCCCTTATGGC	2760	JOURNAL			
D	2813	GGAGATGTTCTTGGTCCCTTCATCAATGCTTACCTCATTTGCGAGCCCTTATGGC	2872	REFERENCE			
Q	2761	TGTGCTGAACAGACATGATAATTTTGTCTCAATATTTACATTTTGGATTATCTGACT	2820	AUTHORS			
D	2873	TGTGCTGAACAGACATGATAATTTTGTCTCAATATTTACATTTTGGATTATCTGACT	2932	TITLE			
Q	2821	AAAAAGAAACAACCTGACAGATTAATTTGAAAGAAAGCTCTTTCATTTATGAGGCAAGT	2880	JOURNAL			
D	2933	AAAAAGAAACAACCTGACAGATTAATTTGAAAGAAAGCTCTTTCATTTATGAGGCAAGT	2992	REFERENCE			
Q	2881	TACAGAGAGAACTTCTTATACAGAGAGAGATGCTCTTTCAGTCTTTGGGAATTTAT	2940	AUTHORS			
D	2993	TACAGAGAGAACTTCTTATACAGAGAGAGATGCTCTTTCAGTCTTTGGGAATTTAT	3052	TITLE			
Q	2941	GACCTTCTGGAGACATTTGTTGCTGAGCTTTGTTTAAAGATTTTCTTGAAGCCGAT	3000	JOURNAL			
D	3053	GACCTTCTGGAGACATTTGTTGCTGAGCTTTGTTTAAAGATTTTCTTGAAGCCGAT	3112	REFERENCE			
Q	3001	CTTACATAGATATTGATCAGAAATGCTTACACAGAAACATACCTGGCTAAAGACAT	3060	AUTHORS			

## FEATURES

source

## CDS

1..4735  
 /organism="Mus musculus"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:10090"  
 /cell\_line="NIH3T3"  
 343..4671  
 /note="GARP"  
 /codon\_start=1  
 /product="GPI-anchored alpha-2 macroglobulin-related protein"  
 /protein\_id="FAM00021.1"

/db\_xref="GI:20070080"

/translation="MRSRLLSAHLICCAVALAAGSRPLVTAPGIIIRGANTYIG  
 VILNENSPQVLYKQVPRKIAASKSKSISLEABGVFRHGFRTLVLPALPLSADKYR  
 LHNIGSENEIVSNKTRFLPESKISISLIVIDKAYKFKQEKVFLVLCSLDKYR  
 TSVDIPKIDPSNVLIQWFSQKGLDVAIVKQSLNSPIFGWMSIQVNDQYQSF  
 QVLEAYLPKFEVVTQPLYSKQKNSVLAQYVGRVKSLSLTFPLPSFMCK  
 KNITKSFBNIGFANFSDNFEKVNPLDVSSEYENDPSPGAEIATATY  
 SLGISRMASNTVPRFKQNDYIIEIPYTLVLRKNSFTATVVRSDGNLPEEEN  
 DLVTVTQRKNHPSQRQENDYIOTVATTIPONKIIEFPWMSISELQKAYFL  
 DGSSVTVSMFTSPSKTYIQKTRREYIKVSPFLMWSGNRQFDSLISWYSKQL  
 VVAGKQSRFTSLPESAWAPACIAIYVIAEGELINDILKIPVLVENVKRLWS  
 KPLVKSIDKYSLEISATQSDSLVGIYAVVKSITLWNSISIMETVHELELVNTY  
 LGMFNPSFVPOECGLMVTLDATLIDNDSIDEVYDREYSEYRFEENANLUPEDSS  
 VNVNVRKNFPEPMITLDAVMSKIYEEFEVTPDSITSWVASAFVISEDIGGLTIV  
 PAELQAFPPFLPLNLPYVINGEERFALVSVNPKDTIKVYLILBESDFPIIMTS  
 NDNGNIYKRTVQVPRDNGVTLPPIKPHLGEIPVTVPASPTASDAVYQIVAPE  
 GLEKSYKSVLIDLIDSNVESKQSWRFSPPTVIGSERVQITAGDILGSSINGLS  
 SIIRMBYGGGEOMNIYFAPNIIYLDVTRKOKLTVLKELALSVMQGYORELLYORE  
 DGSFSAFGDIDSSGTLVSAFVLRGLADYIIDIODVLRHTYTLNMAKKRNGEEM  
 EBRVYHSELQGTSPVTLTAYIVSVYGYKKYQPNIDVOSIKRLEPFRSGIDN  
 YTLAITSVALSTVSGSKAEALNLMQRSEKEDTQFWLSGSPALSGSQWPKVDLEI  
 AAVALAHRTLHVSEGIPIVWRMLIQDRNSLGEVSPQDVLVALKALSEFSAVHKENT  
 DIQITVTPGIPRISIHFRIDSONLFLHQBELHALDPIVNVNSAHSGSPALCOLANDY



NRKSSSRKRRSTENOEVPDI.DVIVNNEDI.SHLNLANCTSHLSERTGMTLMEVNL  
LSGFSASDSIPLSETLKKEYDNKMLYLDVSNESQCVNIPIVRDXYKNINDGS  
VSEMDYPERRQAVRSYNTQVTKSSCYLSPDTRKSHDGTADSLRRSSSLIVFCSVL  
LVEVQH"

BASE COUNT 1338 a 1034 c 1087 g 1276 t

Query Match 63.2%; Score 2709.2; DB 10; Length 4735;  
Best Local Similarity 77.9%; Pred. No. 0;  
Matches 3393; Conservative 0; Mismatches 873; Indels 90; Gaps 7;

QY 1 ATGCAAGGCGCCACCGTCTCTGACCGCGCCACCTCTTGCGTGTGACCGCGCGCTG 60  
DB |||||  
343 ATGCGGAGTCGAGGCTCTCTGAGCGCGCCACCTCTCTGCTGTGCGCGCGCTG 402  
QY 61 GCGGTGGCTCCCGGGGCTGGTTTCTGTGACAGCCCGAGGATCATCGCCCGGAGGA 120  
DB |||||  
403 GCC--GCCCGAGGGTCCCGCTTCTGTGTACAGCCCGAGGATCATCGCCCGAGGA 459  
QY 121 AATGTACTATTGGGGTGTGAGCTTCTGAAACACTGCGCTTCAAGGTGACTGTGAGCG 180  
DB |||||  
460 AATGTACTATTGGGGTGTGATCTCTGAAATAAGCCCCCAGAGTCTTGTAAAGCT 519  
QY 181 GAGCTGCTCAAGACAGATCAAACTCACTGTCTCTGTCTGAGAGCAGAGAGCTTTT 240  
DB |||||  
520 CAGGTGTTCAAGATGACTTCCAACAATGAGATCCATCTGGAGACAGAGAGCTTT 579  
QY 241 GAAAAAGGCTCTTTTAAACACTTACTCTTCACTCACTCACTCTGTAAGAGTGAAGTG 300  
DB |||||  
580 CACAGAGGCGCATTTCAAGACTCTCGTTCTTCGGGCACTCTGTGAGCTGAGATGAAG 639  
QY 301 ATTATGAGCTACGTGTAAACCGGACGTACCCAGAGATGATTTTCTCTAATAGTACC 360  
DB |||||  
640 ATTATGAGCTACACTATAAGGACAAATGAGAAATGATGTATTTCTCAACGAGACA 699  
QY 361 CGCTTATCATTTGAGACCAAGAGATATCTGTCTTCACTTCAACAGAGAGCTTATAC 420  
DB |||||  
700 CGTTTAACTTTGAGGCAAGAGATATCTGTCTTCACTTCAAGAGAGATCTTAC 759  
QY 421 AAGCCAAAGCAAGAGATGAGTTTGGCATTTGACCTTCTCAAGATTTTAAAGCTTAC 480  
DB |||||  
760 AAGCCGAAAGCAAGAGATGAGTTTGGGCTCTCACTGTCTCAAGATTTTAAAGCTTAT 819  
QY 481 AAAACCTCTTAAACATTTCTCATTAAGACCCCAATCAATTTGATCCAGAGTGTG 540  
DB |||||  
820 AAGACATCCGTGACATTTTCAATAAGATCCGAATCTAATGTATCCAAAGTGTG 879  
QY 541 TCAACAACAAGATGATCTGAGATCATTTCCAAAATTTTCAAGTATCTTCCATCCATA 600  
DB |||||  
880 TCTCAAAAAGTATCTGAGATCGTTCCAAAATTTTCAACTATCTCTAATCCATA 939  
QY 601 CTGTGACTGTGCTTATCAAGTTCAGTGAAGTACACAGACATTTATCAATCATTTGAG 660  
DB |||||  
940 TTTGGTACTGTGCTTATCAAGTTCAGTGAAGTACACAGACATTTATCAAGCATTTGAG 999  
QY 661 GTTTCAGATATGATTTTACAAAATTTGAGTACCTTGTGAGACACCATTTATTTGTTCT 720  
DB |||||  
1000 GTCTTGAATATGTGTATCCAAAGTTCGAAGTCAACGTGACAGACCCGCTGTATTTCC 1059  
QY 721 ATGAATTTAAGCATTTAAATGATACATCAAGCAAGATATCATATGGAAGCCAGTG 780  
DB |||||  
1060 TTGAATCTAAACAGTTAAATGATCTCGTCAATGCAAGATATCATATGGAAGCCAGTG 1119  
QY 781 AAGAGGAGATGATCGTTACATTTTATCTTATCTTTTGGGAAAGAAATATTT 840  
DB |||||  
1120 AAGGGAAGTCTGTCACTTATTTTATCTTTGTCTTTTGGGAAAGAAATATATC 1179  
QY 841 ACAAATAAATTAATGATGATCTGCAAACTTCTCTTTTATGATGAAGATGAA 900  
DB |||||  
1180 ACAAATAAATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1239  
QY 901 AATGTAAT-----GATTTCTTAATGATGATTTCTGAATACCTGAT 942  
DB |||||

DB 1240 AAGTAATGAATTTGAAACCGCTTACCGATGTCTTCGAGAGGAGTTATGAGATGTGAC 1299  
QY CTATCTTCCCTGACAGCATGATAATTTTAAACAGTACAGATCATGTTACAGATAT 1002  
DB |||||  
1300 CGGTCTTCTTCTGGGCGGAGATTTATAGTACTGTGAGAGGATCTCAAGAGTATC 1359  
QY 1003 TCAAGAAATGTAAGCATTAATGTGTTCTTCAAGCAACATGATTTATCATGATGATTTT 1062  
DB |||||  
1360 TCAAGAAATGGAAGCAACATGTGTTTTCAGAACAAATGATTTATCATGATGATTTT 1419  
QY 1063 GATTTACTACTGTGTTGAGACCATCTTCACTTCAAGCCATGTTGATGATGATGATGATG 1122  
DB |||||  
1420 GATTTATGAGTGTCTTGAACCAATCTTCACTTCAAGCCATGTTGATGATGATGATGATG 1479  
QY 1123 GCTATGGAACCAACATGATCTTGAAGAAAGAAATTAATGATGATGATGATGATGATG 1182  
DB |||||  
1480 TCTGATGCAATCTAATCTTCTGAGAAATGAAATGATTTGTTGATGATGATGATGATG 1539  
QY 1183 CAGAGAACTATCTGAGTACTGTGAGCGGATCTTAACTGTGAAATCAGAAATGGAAGCT 1242  
DB |||||  
1540 CAGCGAAATTAACCACTGAGAG-----TCAGAGATCAGAAATGATGATTTAT 1590  
QY 1243 GTTCAAAAATTAATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1302  
DB |||||  
1591 ATCCAGACCGTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1650  
QY 1303 CTGAGAGATTCAGTACGATCAAGTGAAGGCTATTTCTTGTGATGATGATGATGATGATG 1362  
DB |||||  
1651 ATGTCATTTCCGGGAGCTACACTGAGGCTTATTTCTTGTGATGATGATGATGATGATG 1710  
QY 1363 GCAGTTCATGATCTGTTTAACTCTCTGATGATGATGATGATGATGATGATGATGATGATG 1422  
DB |||||  
1711 ACAGTCCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1770  
QY 1423 GAAATTAATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1482  
DB |||||  
1771 GAAATTAATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1830  
QY 1483 GATTAATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1542  
DB |||||  
1831 GATTAATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1890  
QY 1543 ACAATGTTCTCTTAAACACAGAAATTTCTTGAATCTCAAAAGCTGTGATGATGATGATG 1602  
DB |||||  
1891 AGAATTTCTCTTAAACACAGAGGCTTCTGAGGCTCAAAAGGCTGTATATGATGATGATG 1950  
QY 1603 TATATTAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1662  
DB |||||  
1951 TATATTAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2010  
QY 1663 TTTTAAATTAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1722  
DB |||||  
2011 TTTGAAATTAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2070  
QY 1723 TCTCTTGAAGTCTGTGACACAGGCTGATCCATATGTTGGATGATGATGATGATGATGATG 1782  
DB |||||  
2071 TCTCTCAAGATCTGTGACACAGGCTGATCCATATGTTGGATGATGATGATGATGATGATG 2130  
QY 1783 AGTGTGAATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1842  
DB |||||  
2131 AGTGTGAATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2190  
QY 1843 GAACTTTTAAACAGAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1902  
DB |||||  
2191 GAACTTTTAAACAGAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2250  
QY 1903 GAACTGTGATCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1962  
DB |||||  
2251 GAACTGTGATCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2310  
QY 1963 TATGACAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2022  
DB |||||  
2311 TACGATTAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2370

QY 2023 ATTGATGACTTTCTTTGGGTAGCAGTCCACATGTCGGAAGCAATTTTCAGAGACTTGG 2082  
 Db 2371 TTTGAGATGCTTTCTTCAAGTATCAATGATCAATGATGAGGAAATTTTCAGAAACATGG 2430  
 QY 2083 ATTTGGCTAGACACCAACATGGGTTACAGGATTTACAGAAATTTGAAGTACTGTAAT 2142  
 Db 2431 ATTTGGCTAGACACCAACATGGGATTTAGATTTATGAAAGAAATTTGAAGTACTGTAAT 2490  
 QY 2143 GATTCTATCACTCTTGGGTGCTACTGTTTGTGATCTGAGGAGCTGGGTCTTGG 2202  
 Db 2491 GACTCTATCACTCTTGGGTGCTCTTGGGTGCTTGGGTGCTTGGGTGCTTGGGTGCTTGG 2550  
 QY 2203 CTAAACACTACTCCAGTGGAGCTCCAGCCTTCCACCACTTTTCAATTTTGTGATCTT 2262  
 Db 2551 TTAAACACCTTTCCAGCAGAGCTGCAAGCTTCCACCACTTTTCTTGGGTGCTTGGGTGCT 2610  
 QY 2263 CCTTACTCTGTTTACAGAGGTGAGAAATTTGCTTGGAAATTAATTAATTAATTTG 2322  
 Db 2611 CCTTACTCTGTTTACAGAGGTGAGAAATTTGCTTGGAAATTAATTAATTAATTTG 2670  
 QY 2323 AAAGATGCGACCTGAGGTAAAGTATCAATGAGAAAGTACCAATTTGATTAATTAATG 2382  
 Db 2671 AAAGATGCGACCTGAGGTAAAGTATCAATGAGAAAGTACCAATTTGATTAATTAATG 2730  
 QY 2383 ACTTCAAGTGAATTAATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2442  
 Db 2731 ACTTCAAGTGAATTAATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2790  
 QY 2443 GGGGCACTGTTCTTTTCCATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2502  
 Db 2791 GGGGCACTGTTCTTTTCCATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2850  
 QY 2503 ACAGCTCTTCAACCACTGTTCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCT 2562  
 Db 2851 ACAGCTCTTCAACCACTGTTCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCT 2910  
 QY 2563 GGAATAGAAATATCATATTCATCAATCTTATTTAGACTTGAATGAGTAAAGCTAAG 2622  
 Db 2911 GGAATAGAAATATCATATTCATCAATCTTATTTAGACTTGAATGAGTAAAGCTAAG 2970  
 QY 2623 AGTACCTGAGAACTTTGAGTTTCTCATTTCTCTCTATACAGTGAAGTGAAGAA 2682  
 Db 2971 AGTACCTGAGAACTTTGAGTTTCTCTCTCTATACAGTGAAGTGAAGAA 3030  
 QY 2683 GTTGAATCACTGAAATTTGAGATGTTCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2742  
 Db 3031 GTTGAATCACTGAAATTTGAGATGTTCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCT 3090  
 QY 2743 ATTGCGATGCTTTATGCTGTGTGTGAGCAGAACTGATTAATTTTGTCTCAATATTTAC 2802  
 Db 3091 ATTGCGATGCTTTATGCTGTGTGTGAGCAGAACTGATTAATTTTGTCTCAATATTTAC 3150  
 QY 2803 ATTTGGATTAATGCTAATAAAGAAACACTGACAGATTAATTTGAAGAAAGCTCTT 2862  
 Db 3151 ATTTGGATTAATGCTAATAAAGAAACACTGACAGATTAATTTGAAGAAAGCTCTT 3210  
 QY 2863 TCATTTATGAGGCAAGTTACAGAGAACTTCTCTATCAGAGGAAAGTGTCTTTC 2922  
 Db 3211 TCATTTATGAGGCAAGTTACAGAGAACTTCTCTATCAGAGGAAAGTGTCTTTC 3270  
 QY 2923 AGTCTTTTGGGAAATTAATGCTTGTGTGTGAGCAGTGTGTGTGTGTGTGTGTGTGTGT 2982  
 Db 3271 AGTCTTTTGGGCAATTAATGCTTGTGTGTGAGCAGTGTGTGTGTGTGTGTGTGTGTGT 3330  
 QY 2983 TGTTCCTTGAAGCGATCTTCAATAGATTAATGATGAGATGTTGTTACAGAAATAC 3042  
 Db 3331 TGTTCCTTGAAGCGATCTTCAATAGATTAATGATGAGATGTTGTTACAGAAATAC 3390  
 QY 3043 ACTTGGCTTAAGAGCAATCAAAATCCAGGCTGAATTTTGGATCCAGGAAAGTGTAT 3102  
 Db 3391 ACTTGGCTTAAGCAATCAAAATCCAGGCTGAATTTTGGATCCAGGAAAGTGTAT 3450

QY 3103 CATATGAGCTTCAAGGTGCGCAATTAAGTCCAGTAACTTACAGCTTAATTTGTAAT 3162  
 Db 3451 CACAGTAACTTCAAGGTGCGCAATTAAGTCCAGTAACTTACAGCTTAATTTGTAAT 3510  
 QY 3163 TCTCTCTGGGATTTAGAAATTAAGTCAAGCTTAATTAAGTCAAGCTTAATTTGTAAT 3222  
 Db 3511 TCTCTCTGGGATTTAGAAATTAAGTCAAGCTTAATTAAGTCAAGCTTAATTTGTAAT 3570  
 QY 3223 TTGAGTCTAAATTAAGTCAAGCTTAATTAAGTCAAGCTTAATTTGTAAT 3282  
 Db 3571 TTGAGTCTAAATTAAGTCAAGCTTAATTAAGTCAAGCTTAATTTGTAAT 3630  
 QY 3283 GCAATGCTAATGAGGAGTCTTAAAGCAAGAACTTTGAATGCTGACTGAG 3342  
 Db 3631 GCAATGCTAATGAGGAGTCTTAAAGCAAGAACTTTGAATGCTGACTGAG 3690  
 QY 3343 GCAGAAACAAAGGTGCAATGCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3402  
 Db 3691 GCAGAAACAAAGGTGCAATGCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3750  
 QY 3403 TGGCAGCAGGCTCTCTGATTAATTAAGTCAAGCTTAATTTGTAAT 3462  
 Db 3751 TGGCAGCAGGCTCTCTGATTAATTAAGTCAAGCTTAATTTGTAAT 3810  
 QY 3463 CAATTTCAAGCTTCTGAGGAGTCTTAAAGTCAAGCTTAATTTGTAAT 3522  
 Db 3811 CAATTTCAAGCTTCTGAGGAGTCTTAAAGTCAAGCTTAATTTGTAAT 3867  
 QY 3523 TTGGGTGCTTTGATTAATTAAGTCAAGCTTAATTTGTAAT 3582  
 Db 3867 TTGGGTGCTTTGATTAATTAAGTCAAGCTTAATTTGTAAT 3927  
 QY 3583 GCAGCCCTTAATTAAGTCAAGCTTAATTTGTAAT 3642  
 Db 3927 GCAGCCCTTAATTAAGTCAAGCTTAATTTGTAAT 3987  
 QY 3643 CCAAGTCTT----- 3651  
 Db 3987 CCAAGTCTT----- 4047  
 QY 3652 CTGCTGT 3711  
 Db 4047 CTGCTGT 4107  
 QY 3712 TGTGCTGCTTAATTAAGTCAAGCTTAATTTGTAAT 3771  
 Db 4107 TGTGCTGCTTAATTAAGTCAAGCTTAATTTGTAAT 4167  
 QY 3772 ATCCAAATCAAGAGCTTTGATTAATTAAGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3831  
 Db 4167 ATCCAAATCAAGAGCTTTGATTAATTAAGTGTGTGTGTGTGTGTGTGTGTGTGTGT 4224  
 QY 3832 AATCATGTGATTTGAATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3891  
 Db 4224 AATCATGTGATTTGAATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 4284  
 QY 3892 CTTATGGAAGTAACTTAATTAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3951  
 Db 4284 CTTATGGAAGTAACTTAATTAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 4344  
 QY 3952 GAGACAGTGAAGAAAGTGAATTAATTAAGTGTGTGTGTGTGTGTGTGTGTGTGTGT 4011  
 Db 4344 GAGACAGTGAAGAAAGTGAATTAATTAAGTGTGTGTGTGTGTGTGTGTGTGTGTGT 4404  
 QY 4012 AATGAAACCAAGTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 4071  
 Db 4404 AATGAAACCAAGTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 4464  
 QY 4072 CAAGAGCTTCAAGT 4131  
 Db 4464 CAAGAGCTTCAAGT 4524  
 QY 4132 TACAACTGTAAGTGAAGT 4191

DB	4525	TACAAACACCCAGGTGAAGCTGTCTCTCAATGTTACCT--CAATCCAGACACCAACTGCAAG	4581
OY	4192	CTTGTGAGATGAGCTTCACAGGCTCCCATCATCTTTCCTGT	4251
DB	4582	TCACACACGAGCGAGCCACGACCTCCCTTCGACGCTTCCTGTCTTGTGC	4641
OY	4252	TTCAAGCTCTGTACTTATGAGACTTTGGCTGCA	4287
DB	4642	TCCTGCTCTGTACTTGTCTTCACATGATTATTA	4677
RESULT 11			
LOCUS	BC052443	5644 bp	linear
DEFINITION	Mus musculus cDNA clone MGC:63463 IMAGE:6406841, complete cds.		
ACCESSION	BC052443		
VERSION	BC052443.1	GI:30851465	
KEYWORDS	MGC.		
SOURCE	Mus musculus (house mouse)		
ORGANISM	Eumetazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.		
REFERENCE	Strauberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G., Klausner, R.D., Collins, F.S., Wagner, L., Shennan, C.M., Schuler, G.D., Altschul, S.F., Zeeberg, B., Bietow, K.H., Schaefer, C.F., Bhat, N.K., Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Heien, F., Ditchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L., Steadton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L., Schreier, T.E., Brownstein, M.J., Udell, T.B., Toshiyuki, S., Carninci, P., Prange, C., Rana, S.S., Loquellano, N.A., Peters, G.J., Abramson, R.D., Mullaly, S.J., Bosak, S.A., Moban, P.J., McErmann, K.D., Malek, J.A., Gunaratne, P.H., Richards, S., Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W., Viall, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fahey, J., Helton, E., Kettelman, M., Madan, A., Rodriguez, S., Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y., Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butlerfield, Y.S., Krzywinski, M.I., Skalka, U., Smal, D.E., Scherch, A., Schein, J.E., Jones, S.J. and Marra, M.A., Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences		
TITLE	human and mouse cDNA sequences		
JOURNAL	Proc. Natl. Acad. Sci. U.S.A.	99 (26), 16899-16903	(2002)
MEDLINE	22388257		
PUBMED	12477932		
REFERENCE	2 (bases 1 to 5644)		
AUTHORS	Strauberg, R.		
TITLE	Direct Submission		
JOURNAL	Submitted (15-MAY-2003) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA		
REMARK	NIH-MGC Project URL: <a href="http://mgc.nci.nih.gov">http://mgc.nci.nih.gov</a>		
COMMENT	Contact: MGC help desk Email: <a href="mailto:cgapbs-remail.nih.gov">cgapbs-remail.nih.gov</a> Tissue Procurement: Dr. Jim Ijdo, University of Iowa cDNA Library Preparation: M. Bento Soares, University of Iowa cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN) DNA Sequencing by: University of Iowa, Dr. M. Bento Soares and Dr. Thomas L. Casavant. Web site: <a href="http://genome.uiowa.edu">http://genome.uiowa.edu</a> Contact: <a href="mailto:bento-soares@uiowa.edu">bento-soares@uiowa.edu</a> ; <a href="mailto:tom-casavant@uiowa.edu">tom-casavant@uiowa.edu</a> Bonaldo, M.F., Akabogu, I., Bair, T., Bair, J., Crouch, K., Davis, A., Fisher, K., Keppel, C., Kucaba, T., Lebeck, M., Melo, A., Schaefer, K., Schreier, T., Smith, C., Snir, E., Tack, D., Trout, K., Walters, J., Casavant, T., Soares, M.B.		

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNLN at: <http://image.lnl.gov>

Series: Plate: Row: Column: 0.

Location/Qualifiers

```

SOURCE
1. 5644
    /organism="Mus musculus"
    /mol_type="mRNA"
    /strain="C57BL/6"
    /db_xref="taxon:10090"
    /clone="MGC:63463 IMAGE:6406841"
    /tissue_type="Brain enriched mouse brain 12.5dp"
    /clone_id="N1H_BMAP_F00"
    /lab_host="DH10B"
    /note="vector: pYX-ASC"
    105..4433
    /codon_start=1
    /product="Unknown (protein for MGC:63463)"
    /protein_id="AAH52443.1"
    /db_xref="GI:30851466"
    /translation="MRSRRLLISAHLCLCAVALAAPGRSFLVTAPEGIRPAGNVITG
VDLLENSPQVILVAQVFKIASNKSRSIIAEAGFPHGHRKTLVLPALPLISADKIYEE
LHNGENSENIEFSPNRRLTPRESKISVLQTPKAPKPKQYKFRLLTICSLDKPRR
TSVDIFIKDKPSNVIOQMFESQKDLGVSKIFQKLSNPIDGDSIOVQVNDQYQSF
QVLELYLPKFEVIVQTEPLCYSLSKQSLNGSVIAKYTGKRVKSLSTLPLSPWKGK
KLTGKSPFENGAFNFSFDNYEMKKVMKLKELTVSDESNVWVSPFGPAEIIATVEE
SLNIGSMASTNVFQKHQYIIEIPDTVTLKSLNATVAKVSRSGNDQTPREINEN
DLVATVTRKNNHESRODEMIOYIOWVNTIPONGIKIEPMPNISGELQAKAYPL
DGSVTVTHSMFTSPKTYIQLKTRDIYIVGSPFDLMVSGNRPKDSITVNVISLQK
VVAQKOSRTFSLTPESAWAPKACIIAYIAEAGEIINDILKILPVQVFNKELIFPMS
KPIVPSKPDKXSRISATQSDLVGIVANDSVTLMEWNSIIMETWHELELVNTEYY
LGMHMFAYFOECGLVLTDLIRDSIDEVYDTEYSERFAENANLVDPEDASS
VNNVHVKNNPEPTIMLDAYMGKIIYEEFNVYDSITSMVAAFVISEDGFGELTVTS
PAELHAFQPEPLFLNLPYSVIRGEFALVSIYVNLKDTIKVILLIESSPDILMTS
NDTNGTITRKTVQVPRDNGTVLVPKIKPHHGEIPITVTAASPASAVTOTIYKKEE
GIEKISKSVLVDLDSNVESKQDSMFSPPTLVISERQVLTADILGSSNGHS
SLIRMPYGCGBORNIYPAPIYILDITKQKRLNINKEKALSMRGQYORIELYQRE
DEGSFAGSDIDSSGTSWLSAFVLCRFCEADYIIDVDLHRTYTMANHKFNGEPMW
EBSGVHSELDGGTSPVTLTAIVTSLVYKKYQIPDVODSIKFLPEPERSGISDN
YTLAISYALSTYSGPKAEELNLMORSKEGDTOPWLSGPAISGMSORASYDIET
AAYRLALHTLHHVSEGIPIVWRMLIQDNISIGCFSTQDITVALKALSEFSLVHKENT
DIQULTVGPQIPRSIHRIDSONFLHQEELHLADEITYVNSAHSGSFALQQLNVY
NNGSGSSKRRSRSTENQEDVDLYIVNNEIDI SHLNTQCTSHLGSRTGQVLENNI
LVGSSASDSIPISETLKVEYDNGKNILVDSVNESQVCNIIPTVADYKSNLRDSS
VSWMDYPERRAQVRSVNTQVTKLSSCLSPDNTCKSHTDATSLRRSSSLVPCSVL
LYVQHY"

BASE COUNT      1556 a      1235 c      1251 g      1602 t
ORIGIN
Query Match      63.2%; Score 2709.2; DB 10; Length 5644;
Best Local Similarity 77.9%; Pred. No. 0;
Matches 3393; Conservative 0; Mismatches 873; Indels 90; Gaps 7;

OY      1 ATGCAAGGAGCCACCGCTCTGTGACCGCGCCCTACTCTCTGTGCTGTCACCGCGCGCTG 60
Db      105 ATGGCAGTGGAGAGGCTCTGTGAGCGCGGCCACCTTCTGTGCTTGCGCGGTGCGCTG 164

OY      61 GCCGTGCTCCCGGCGCTCGGTTTCTGTGTGACAGAGCCCAAGGATCATCAGGCCCGAGCA 120
Db      165 GCC---GCCCAAGGCTCCGCTTCTGTGTGACAGAGCCCAAGGATCATCAGAGCCGAGCA 221

OY      121 AATGTGACTATGGGGGTGAGGCTTCTGGAACACGTGCCCTTCAACAGGTGACGTGTAAGGCG 180
Db      222 AATGTGACTATGGGGGTGAGATCTTCTGTGAATAATGCCCCCACAAGGTCCTTGTGTAAGGCT 281

OY      181 GAGCTGCTCAAGACAGATCAAACTTCACTGTCTGTCTGCTGTGAGAGCAAGAGAGTCTTT 240
Db      282 CAGGTGTTCAAGATAGTCTTCCAACAATTCAGATCCATCTCTGGAAGCAAGAGAGTCTTT 341

OY      241 GAAAAAGGCTCTTTAAGACACTTACTTTCATCACTACTCTTGAACAGTGCAGATGAG 300
Db      342 CACAGAGGCAATTCAGACTCTGTCGTTCTTCGCGCACTCACTCTGACAGTGCAGATAG 401

OY      301 ATTATAGGCTAGCGTGAACCGGACGTAAACCAGAGATGATTTTATCTCTAATATGATACC 360
Db      402 ATTATAGGCTACATTAACCGGACATCAAGAGATGAGATCTGATTTCTTCACACAGGACA 461

OY      361 CGCTTATCATTTTGAGCAAGAGAAATCTGTCTTCAATTCATAACAGACAGAGCCTTTATC 420

```

462 CGTTTAACTTTGAGAGAGAGATATCTGCTCCATTCAGAGAGAGAGCAATTCAC 521  
421 AAGCCAAAGCAAGAGTGAAGTTTGCATTTGACATCTTCTGAGATTTTGAAGCTTAC 480  
522 AAGCCGAAGCAAGAGTGAAGTTTGCATTTGACATCTTCTGAGATTTTGAAGCTTAT 581  
481 AAAACCTCTTTAAACATTTCTGATTAAGGACCCCAATCAATTTGATCCAAAGTGTG 540  
582 AGGACATCCGTGACATTTTCATTAAGATCCGAAGTCTAATGATCCAAAGTGTG 641  
541 TCACAAACAAAGTGAATCTTGAGATCATTTCCAAACCTTTGAGCTATCTTCCATCCATA 600  
642 TCTCAGAAAGGTGATCTTGAGATCGTTTCCAAACCTTTTCAACATCTCTCAATCCATA 701  
601 CTGAGTGCATGCTCATTCAGATTCAGATGATGACGAGATCATTAATCATCATTCAG 660  
702 TTTGATGATGCTGATTCAGATTCAGATGATGATGATGATGATGATGATGATGATGAT 761  
661 GTTTCAGAAATATGATTAACCAAAATTTGAGTGAATTTGACAGACATTAATTTGCT 720  
762 GTTTCAGAAATATGATTAACCAAAATTTGAGTGAATTTGACAGACATTAATTTGCT 821  
721 ATGAATTTAAGCATTTAATGATGATGATGATGATGATGATGATGATGATGATGAT 780  
822 TTGAAGTCTAAACAGTAAATGAGCTCCGTCATAGCAAGATTAATGATGATGATGAT 881  
781 AAAGGAGAGTGAAGTGAATTTTACCTTATCTTTGAGGAGAGAGAGAGAGAGATTT 840  
882 AAGGAGAGTGAAGTGAATTTTACCTTATCTTTGAGGAGAGAGAGAGAGAGATTT 941  
841 ACAAACCATTTAAGATTAATGATGATGATGATGATGATGATGATGATGATGATGAT 900  
942 ACAAACCATTTAAGATTAATGATGATGATGATGATGATGATGATGATGATGATGAT 1001  
901 AATGATAT-----GATTTCTTCAATGAGCTTTGATGATGATGATGATGAT 942  
1002 AAGGATATGATTTGAAACCGCTAAGAGATGCTGAGAGAGAGATTAAGAGAGATGAT 1061  
943 CTATCTTCCCTGAGAGAGATTAATTTTACCAAGAGAGATGATGATGATGATGAT 1002  
1062 CCGTCTTCTGAGAGAGATTAATTTTACCAAGAGAGATGATGATGATGATGATGAT 1121  
1003 TCAAGAAATGATGATTAATGATGATGATGATGATGATGATGATGATGATGATGAT 1062  
1122 TCAAGAAATGATGATTAATGATGATGATGATGATGATGATGATGATGATGATGAT 1181  
1063 GATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1122  
1182 GATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1241  
1123 GCTAATGAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1182  
1242 TCTAATGAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1301  
1183 CAGAGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1242  
1302 CAGAGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1352  
1243 GTTTCAGAAATTAATTAATGATGATGATGATGATGATGATGATGATGATGATGAT 1302  
1353 ATTCAGAGAGATTAATTAATGATGATGATGATGATGATGATGATGATGATGATGAT 1412  
1303 CTGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1362  
1413 ATTCAGAGAGATTAATTAATGATGATGATGATGATGATGATGATGATGATGATGAT 1472  
1363 GCAGTTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1422  
1473 ACAGTTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1532  
1423 GAAATATTAAGAGTGAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1482

1533 GAAATATTAAGAGTGGGCTACCTTTTGAATTTGATGATGATGATGATGATGATGAT 1592  
1483 GAGTTAAGTATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1542  
1593 GACTTAAGTATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1652  
1543 ACAATGTTCTTTTAAACACAGAGAAATTTTGAATCTTCAAAAGCTGATGATGATGAT 1602  
1653 AGAATCTTCTTTTAAACACAGAGAGCTTCTGAGCTTCAAAAGCTGATGATGATGAT 1712  
1603 TATATTAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1662  
1713 TATATTAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1772  
1663 TTTAAATTAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1722  
1773 TTTAAATTAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1832  
1723 TCTTTAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1782  
1833 TCTCTAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1892  
1783 AGTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1842  
1893 AGTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1952  
1843 GAACTTATTAACACAGAGATTAATTTAAGGATGATGATGATGATGATGATGATGAT 1902  
1953 GAACTTATTAACACAGAGATTAATTTAAGGATGATGATGATGATGATGATGATGAT 2012  
1903 GAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1962  
2013 GAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2072  
1963 TATGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2022  
2073 TATGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2132  
2023 ATTCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2082  
2133 TTTGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2192  
2083 ATTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2142  
2193 ATTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2252  
2143 GATTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2202  
2253 GATTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2312  
2203 CTAAACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2262  
2313 TTTAAACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2372  
2263 CCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2322  
2373 CCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2432  
2323 AAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2382  
2433 AAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2492  
2383 ACTTAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2442  
2493 ACTTAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2552  
2443 GGGGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2502  
2553 GGGGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2612  
2503 ACAGTCTTTTACCACTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2562  
2613 ACAGTCTTTTACCACTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2672

QY	2563	GGAAATAGAAAATCAATATTCACATCCATCTTAATAGCTTGACTGCACATAGGCTACAG	2622
Dp	2673	GGGATAGAAAAGTCGATTCGCAAAATCTGCTCTTAATGGATCTGCACAGATAGCAAGTACGA	2732
QY	2623	AGTACCCCTGAAAACCTTGAAGTTTCTCAATTCCTCTCTATACAGTGAAGCTGCAGTGAAGA	2682
Dp	2733	AGTAGAGCAGCAATCTATAGAGATTCCTCTTCCTCTGTAACGGCATTTGGCAGTAGAAGA	2792
QY	2683	GTTTCAGATCACTGCATTTGGAGATGTTCTTGGTCCCTTCATCAATGGCTTAAGCTCATTTG	2742
Dp	2793	GTTTCAGATCACTGCATTTGGAGATGTTCTTGGTCTTCCTCATCAATGGCTTAATCTTCACTG	2852
QY	2743	ATTGCGATGACCTTAATGGCTGATGGAGACAGAACATGATTAATTTTGGCTCCCAATATTTAC	2802
Dp	2853	ATCCGAGATGCTTAATGGATGATGGAGAACAGAACATGATTAATTTTGGCTCCCAATATTTAC	2912
QY	2803	ATTTTGGATTAATCTGACTTAATAAAGAAAACAATCTGACAGATATTTTGAAGAAAAGAGCTT	2862
Dp	2913	ATTTTGGATTAATCTGACTTAATAAAGAAAACAATCTGACAGATATTTTGAAGAAAAGAGCTT	2972
QY	2863	TCATTTATAGGCAAGGTTACCAAGAGAACTTCTATTCAGAGGAAAGATGCTCTTTC	2922
Dp	2973	TCATTTATAGGCAAGGTTACCAAGAGAGCTTCTATTCAGAGGAAAGATGCTCTTTC	3032
QY	2923	AGTCCCTTTTGGGAATATAGACCCCTCTTGGAGACACTGGTGTGACGCTTTTGTTTTAA	2982
Dp	3033	AGTCCCTTTTGGGACATTTGACTCTTCTTGGAGACACTGGCGTGCAGATTTGTTTAA	3092
QY	2983	TGTTTCCCTTGAAGCCGATCCTTAATACATGATATTAATACAGATGATTAACAGAAACATAC	3042
Dp	3093	TGCTTTCTGGAAAGCTGATTAATTAATAGATTAATTAATCAGATGATGTTTACACAGAACTAT	3152
QY	3043	ACTTGGCTTAAAGGACATCAGAAATCCACCGGTGAATTTTGGGATTCACAGAAAGATGAT	3102
Dp	3153	ACTTGGCTTAAATGACATTAAGAAATTCATGATGATTTTGGAGCCACAGAAAGATGAT	3212
QY	3103	CATATGAGCTTCAAAGGTGGCAATPAAAAGTCCAGTAACCTTACAGCCTATATTTGTAAT	3162
Dp	3213	CACAGTGAATCTTCAAAGGTGGCAACCAAAAGCCCAATTAACATTAACGCTATATTTGTAAT	3272
QY	3163	TCTCTCCCTGGGATATPAAAAGTATACAGCTTAACTTGAATGGCAAGATCTATTCATTTT	3222
Dp	3273	TCTCTCCCTGGGATPAAAAGTATACAGCTTAAATTCATGTAACAAAGCTCAATPAAATTT	3332
QY	3223	TTGAGATCTGAATTCAGTAGAGAAATTTACACAAATTAATCTTAGCCCTTAATCTTAT	3282
Dp	3333	TTGGAATTTGAATTCAGCAGAGGAATTTCCGACATTAATTAATCTTGAATTAATTCCTAT	3392
QY	3283	GCAATTTGTCATCAGTGGGGAGTCTTAAAGCGAAGAAAGCTTTGGAATAATGCTACTTGGAGA	3342
Dp	3393	GCCTGTGTCACAGTCCGGAGGCCCTTAAGCAGAAGGAGCTTTGAACTTGTGATGTCACAGA	3452
QY	3343	GCAAGAACAGAAAGTGGGATGCAATTTCTGGGTGTCATCAGATCCAAATCTTCTGACTCC	3402
Dp	3453	TCAGAAAAGGAGAGACACACAGTTCGTGTTGTCATCAGAGACCTGCGGCTCTCGGGGTCC	3512
QY	3403	TGGCAGCAGGCTCCCGATTAATTAAGATTCACAGCTATGCAATGACATGCACTTCTCA	3462
Dp	3513	TGGCAGCAGGCTCCCGATGATGAATTCGCTGTTATGCACTGCTGGGCAACATCTGTG	3572
QY	3463	CAATTTCAAGCTTCTGAGGGAATCCCAATTTAGAGTGGCTTAACAGGCAAAAGAAATAGC	3522
Dp	3573	CA-----CAATGATGTCAGGGAATCCCGGTTATAGAGTGGCTCATTCAGCAAAAGAAACAGC	3629
QY	3523	TTGGGTGTTTTCATTAATCAGAGATTCACATGCTGGCTTTAAAGGCTCTGTCTGAATTT	3582
Dp	3630	CTGGGAGGTTTTCATCCTCAGAGATTCCTGTTGGGCCCTGAAGGCAATTAATCGAATTT	3689
QY	3583	GCAGGCCCAATGAATACAGAAAGCAAAATATCCAAATGACCGGTGACGGGGCCCTAGCTCA	3642
Dp	3690	TCAGGCCCTTGTGCACAGAAAATTCAGATATTCACATGACCTGTCAACAGGGGCCCGGATTC	3749

Oy	3643	CCAGTCTCT-----	CCAGTCTCT-----	3651
Db	3750	CCCGAGATTCATACACTTCCGAATTGACTCCCGAAGCTCTTTCTCTTACACAGAGAG	3809	
Oy	3652	CTTCCTGTGTGATACCGCAATGGCAGTTAATAATTTCCGCAATGGTTTTGGATTGTCTATT	3711	
Db	3810	CTTATGATGGCTAGATATCCCATCAGCGTTTAACTTTCTGTGCATGGCTGGGATTTGTCTATT	3865	
Oy	3712	TGTCACTCAATGTTGTATATATATATGTAAGGCTTCTGGGCTTTCTAGAAACGAATCT	3771	
Db	3870	TGCCAGCTTAATGTTGACTAATAAGTAAAGGTTCAGGTTCTTTAAAGAAGCGGAGATCT	3922	
Oy	3772	ATCCAAATCAAGAAGCCTTGATTGATTTGATGCTGTCTGTAAAGAAATAAAGATGATCTC	3833	
Db	3930	ACCGAAATCAAGAAGTTTGTGATTATGACGTCAATTGTGA---ATTAAGAGAGACGACATT	3986	
Oy	3832	AATCATGTGATTTGAATGATGTGTATACAGCTTTTCGGGCGCGGGATGAGATGGCATGCT	3891	
Db	3987	AGTCACTGAAATCTGAATGTGTGTGACAAAGTCACTTGGTTTACAGAGAGACAGCATGGTTC	4048	
Oy	3892	CTTATGGAAGTTAACTTAATTAAGTGGCTTTATATGAGCTTTCAGAAAGCAATTTCTGTAGC	3951	
Db	4047	CTTATGGAAGTGAACCTTTCTCAGTGGCTTTAGTGATCTTCAGATTTCAATTTCTGTAGT	4106	
Oy	3952	GAGCAGTGAAGAAATGTGAATATGATCATGGAATACTCAACTCTATATTAGATTCTGTA	4011	
Db	4107	GAGACCCCTGAAGAAATGTGAATATGATPACGGGAAACTCAACTTAAATTATGATTTCTGTA	4166	
Oy	4012	AATGAATCCCGATTTGTGTTAATATATTCCTGTGTGAGAACTTTAAAGTTTCAAAATACC	4071	
Db	4167	AATGAATCCCGATTTGTGTTAATATATTCCTGTGTGAGACCTTAAAGTTTCAAAATATT	4226	
Oy	4072	CAAGATGCTTCAAGTGCATATGTGTGATTAATATATGAGCAAGAGACAGGCGGTGAGAAGT	4133	
Db	4227	CGAATGTCTCAATATCTGTATATGACTACTATATGAACCAAGAGACAGGCAAGTGTGGAAGC	4288	
Oy	4132	TACAACCTGTGAAGTGAAGCTGTCTCTCTGTGACCTTTGTCACTGATGTATCCAGGCGTGCCT	4191	
Db	4287	TACAACACCCAGGTGAAGCTGTCTCTCATGTTAACCT---CATCCAGACACCAACTCGCAAG	4343	
Oy	4192	CTTTGTGAGATGAGCTTCAAGGCTCCCATCATCACTCTTCAATGTCATTTTATTTTCTGT	4251	
Db	4344	TCACACACGAGCGGAGCCACGGACTCCCTTGAGATTTCTTGAGACCTTCTGTCTTTTGC	4403	
Oy	4252	TTCAAGCTTCTGTACTTATGGAATCTTTGGGTGGA	4287	
Db	4404	TCGGTCTTCTGTACTTGTGCAACATGTGATTAAGA	4439	

RESULT	12
HSM805581	
LOCUS	HSN805581
DEFINITION	Homo sapiens mRNA; cDNA DKFPZ762L1111 (from clone DKFPZ762L1111).
ACCESSION	AL334478
VERSION	AL334478.1 GI:21740252
KEYWORDS	.
SOURCE	Homo sapiens
ORGANISM	Homo sapiens (human)
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo. 1 (bases 1 to 2938) Blum,H., Baurerachs,S., Mewes,H.W., Weill,B. and Wiemann,S. Direct Submission Submitted (09-JUL-2002) 1. D-85764 Neuberg, GERMANY Clone from S. Wiemann, Molecular Genome Analysis, German Cancer Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de; Munich/Germany) within the cDNA sequencing consortium of the German Genome Project.
TITLE	This clone (DKFPZ762L1111) is available at the RZPD in Berlin. Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059 Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de Further information about the clone and the sequencing project is available
JOURNAL	
COMMENT	



at <http://mips.gsf.de/proj/cDNA/>.  
FEATURES  
source  
1. .2938  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/map="15"  
/clone="DKFZp762L1111"  
/tissue\_type="melanoma (Melo cell line)"  
/clone\_lib="762 (synonym: hmel2). Vector pSPORT1; host DH10B; sites NotI + SalI"  
/dev\_stage="adult"  
1. .2938  
/gene="DKFZp762L1111"  
/gene="DKFZp762L1111"  
/note="similarity to alpha-2-macroglobulin (Limulus sp.)"  
/codon\_start=1  
/product="hypothetical protein"  
/protein\_id="CAD9137.1"  
/db\_xref="GI:21740253"  
/translation="VDSKSVNLMNSNDITMENVVHELELYNTGYLGMFNNSPAVPOE  
CGMLVTDANITKDYIDGVYDNEAYARPMENEGHVIDHDSLSPPHYRKRPEET  
WMLDWMGSKITQEFVETVDSITSWATGFEVISEDGLGLTTPVELDAFPQFPLF  
LNLPSYVIRGEFALBITFENYLDKATEVKVILEKSDFILMTSSINATGHQTLT  
VPSEDAIVLFPPIPTLGEIPITVALSPASDAITOMILVKKEGSEKYSOILLD  
LTDNRLOSTLKSSEFPNVTGSEVQITAGIDVQVQRELLIYQEDSFEAPRGVDS  
MINPAPNITLIDYTKKOLTDNKEKALSPMGQVROGKRELLIYQEDSFEAPRGVDS  
GSTSLAPVLRCELEADPYIDIDQVILHRTYTLKQKNGEWMGRVTHSLQSG  
NKSPTVLTAIVLISLIGYRKOPNIDVQESLHFLSEFSRGIISNVTALITVALSGV  
GSPKAKALNMLTWRAOEGQMOPVWSESLSDSQPRSDIENVAVALSHLQPO  
TSEGIPIKRWLRSORNSLGFASOTDTVALKALSEPALMTERTNIOYLVGSPSP  
SPVKEILDTHNRLLQTAELAVQPTAVANSANGEPALCOLAVVNVKXSSRRR  
SIONOEA.PDLIDVAIKENKDLNHYDLNVCFSFSGRSGMALMEVNLISPMVSEAI  
SLSTETKKEVEIDHKLANLYLDSVNETQPCNIPVNRPKNSNTODASVSTDIYEPR  
QAVRSYSEVLSCLSDVQGRPEDGASGSHHSVLFICFPLLYFMEIWL"  
2843. .2848  
polya\_signal  
/gene="DKFZp762L1111"  
2859  
polya\_site  
/gene="DKFZp762L1111"  
BASE COUNT 887 a 529 c 620 g 902 t  
ORIGIN  
Query Match 57.1%; Score 2448.2; DB 9; Length 2938;  
Best Local Similarity 97.9%; Pred. No. 0;  
Matches 2511; Conservative 0; Mismatches 3; Indels 51; Gaps 1;  
QY 1774 GTTGACAAAAGTGAATCTGATGAATGCTTAATGATATTACATGGAATGGGTC 1833  
Db 1 GTTGACAAAAGTGAATCTGATGAATGCTTAATGATATTACATGGAATGGGTC 60  
QY 1834 CATAGCTTGAACCTTTATTAACAAGATATTATTAAGCATGTTCAAGAACTCTTTGCA 1893  
Db 61 CATAGCTTGAACCTTTATTAACAAGATATTATTAAGCATGTTCAAGAACTCTTTGCA 120  
QY 1894 GTCTTTGAGAATGGAAGCTCTGGATATTGACAGTCAAACTCAGCAAGATTAATT 1953  
Db 121 GTCTTTGAGAATGGAAGCTCTGGATATTGACAGTCAAACTCAGCAAGATTAATT 180  
QY 1954 GATGCTTTATGACAAATGACAAATATGCTGAGAGTTTATGAGAAATGAAGACAT 2013  
Db 181 GATGCTTTATGACAAATGACAAATATGCTGAGAGTTTATGAGAAATGAAGACAT 240  
QY 2014 ATTGTAGATATTCAATGATCTTTCTTTGGGTGACAGTCAATGTCGGAAGACTTTTCCA 2073  
Db 241 ATTGTAGATATTCAATGATCTTTCTTTGGGTGACAGTCAATGTCGGAAGACTTTTCCA 300  
QY 2074 GAGACTTGATTTGGCTGACCAACATGGGTTACAGGATTTTCAAGAAATTTGAAGTA 2133  
Db 301 GAGACTTGATTTGGCTGACCAACATGGGTTTCCAGGATTTTCAAGAAATTTGAAGTA 360  
QY 2134 ACTGTACCTGATTTATCACTCTTGGGTGCTACTGTTTGTGATCTGAGACCTG 2193

Db 361 ACTGTACCTGATTTATCACTCTTGGGTGCTACTGTTTGTGATCTGAGACCTG 420  
QY 2194 GGTCTTGACATAACAATACTACTCCAGTGAAGCTCCAGGCTTCCACCAATTTTCAATTTT 2253  
Db 421 GGTCTTGACATAACAATACTACTCCAGTGAAGCTCCAGGCTTCCACCAATTTTCAATTTT 480  
QY 2254 TTGAATCTTCCCTACTCTGTATTACAGAGTGAAGAAATTTGCTTTGAAATTAATTAATTC 2313  
Db 481 TTGAATCTTCCCTACTCTGTATTACAGAGTGAAGAAATTTGCTTTGAAATTAATTAATTC 540  
QY 2314 AATTATTGGAAGAATGCCACTGAGGTTAAGATTAATCAATGAGAAATGCAAAATTTGAT 2373  
Db 541 AATTATTGGAAGAATGCCACTGAGGTTAAGATTAATCAATGAGAAATGCAAAATTTGAT 600  
QY 2374 AATTCTAATGACTTCAAGTGAATAATAATGACAGGCAACAGAGACCTTCTGTTCCC 2433  
Db 601 AATTCTAATGACTTCAAGTGAATAATAATGACAGGCAACAGAGACCTTCTGTTCCC 660  
QY 2434 AGTAGAGATGGGCAACTGTTCTTTTCCCATGAGGCAACACATCTGGAGAAATTCCT 2493  
Db 661 AGTAGAGATGGGCAACTGTTCTTTTCCCATGAGGCAACACATCTGGAGAAATTCCT 720  
QY 2494 ATCAGACTCAGACTCTTTCAACCACTGCTTGATGCTGTCAACCAAGATGATTTAGTA 2553  
Db 721 ATCAGACTCAGACTCTTTCAACCACTGCTTGATGCTGTCAACCAAGATGATTTAGTA 780  
QY 2554 AAGCGTAAGAGATAGAAAATCATATATTCATATTCATCTTATTAAGCTTGAATGACAT 2613  
Db 781 AAGCGTAAGAGATAGAAAATCATATATTCATATTCATCTTATTAAGCTTGAATGACAT 840  
QY 2614 AAGCTACAGAGTACCTGTAACCTTTGAGTTTCTCATTTCTCTTAATACAGTACTGGC 2673  
Db 841 AAGCTACAGAGTACCTGTAACCTTTGAGTTTCTCATTTCTCTTAATACAGTACTGGC 900  
QY 2674 AGTGAAGAAGTCAATCATCTGCAATTTGAGATGTTCTTGCTCTTCAATGATGCTTA 2733  
Db 901 AGTGAAGAAGTCAATCATCTGCAATTTGAGATGTTCTTGCTCTTCAATGATGCTTA 960  
QY 2734 GCCCTATGATTTGGAGATGCTTATGAGCTGTGAGTGAAGAAACATGATTAATTTGCTCCA 2793  
Db 961 GCCCTATGATTTGGAGATGCTTATGAGCTGTGAGTGAAGAAACATGATTAATTTGCTCCA 1020  
QY 2794 AATATTACATTTTGGATTTATCTGACTAAAGAAACAACTGACAGATTAATTTGAAAGAA 2853  
Db 1021 AATATTACATTTTGGATTTATCTGACTAAAGAAACAACTGACAGATTAATTTGAAAGAA 1080  
QY 2854 AAGCTCTTCAATTTATGAGGCAAGTTTACAGAGAACTTCTATTCAGAGGAAAT 2913  
Db 1081 AAGCTCTTCAATTTATGAGGCAAGTTTACAGAGAACTTCTATTCAGAGGAAAT 1140  
QY 2914 GGCCTTTGAGTCTTTTGGGAATTTATGACCTTCTGGAGCACTTGCTGAGCTTT 2973  
Db 1141 GGCCTTTGAGTCTTTTGGGAATTTATGACCTTCTGGAGCACTTGCTGAGCTTT 1200  
QY 2974 GTTTAAGATGTTTCTTGAAGCGATCTTAATGATATTATGATCAAGATGTTTACAC 3033  
Db 1201 GTTTAAGATGTTTCTTGAAGCGATCTTAATGATATTATGATCAAGATGTTTACAC 1260  
QY 3034 AGAATATCACTTGCTTAAAGCATCAGAAATCCACGCTGAATTTTGGATCCAGGA 3093  
Db 1261 AGAATATCACTTGCTTAAAGCATCAGAAATCCACGCTGAATTTTGGATCCAGGA 1320  
QY 3094 AGAGTATTCATATGAGCTTCAAGGTCGAATTAAGTCAATTAACCTTAAGGCTAT 3153  
Db 1321 AGAGTATTCATATGAGCTTCAAGGTCGAATTAAGTCAATTAACCTTAAGGCTAT 1380  
QY 3154 ATTGTAACTTCTCTCTGAGATATAGAAATATCAGCTTAACATTTGATGCAAGTCT 3213  
Db 1381 ATTGTAACTTCTCTCTGAGATATAGAAATATCAGCTTAACATTTGATGCAAGTCT 1440  
QY 3214 ATCCATTTTGGAGTCAATTCAGTAGAGAAATTCAGACAAATTAATTAATCTAGCCCTT 3273  
Db 1441 ATCCATTTTGGAGTCAATTCAGTAGAGAAATTCAGACAAATTAATTAATCTAGCCCTT 1500



```

OY 3274 ATAACTTATGCAATGTCATGAGTGGAGTCTTAAAGCAAGAGCTTTGAATATGCTG 3333
    |||||
DB 1501 ATAACTTATGCAATGTCATGAGTGGAGTCTTAAAGCAAGAGCTTTGAATATGCTG 1560
OY 3334 ACTTGGAGGCAAGCAAGAAGTGGCAATTTGGTGTGATCAAGATCCAACTT 3393
    |||||
DB 1561 ACTTGGAGGCAAGCAAGAAGTGGCAATTTGGTGTGATCAAGATCCAACTT 1620
OY 3394 TCTGACTCTGGCAGCAGCTCCCTGGATATTAAGTTCAGACCTATGCACTGCTCTCA 3453
    |||||
DB 1621 TCTGACTCTGGCAGCAGCTCCCTGGATATTAAGTTCAGACCTATGCACTGCTCTCA 1680
OY 3454 CACTTCTTAACAATTTGAGCTTCTGAGGGAATCCCAATTATGAGGTGGCTTAAGCAGCA 3513
    |||||
DB 1681 CACTTCTTAACAATTTGAGCTTCTGAGGGAATCCCAATTATGAGGTGGCTTAAGCAGCA 1740
OY 3514 AAAAAATAGCTTGGGTGTTTGTGATCTACTAGAGATACCACTGTGGCTTTAAAGGCTCTG 3573
    |||||
DB 1741 AAAAAATAGCTTGGGTGTTTGTGATCTACTAGAGATACCACTGTGGCTTTAAAGGCTCTG 1800
OY 3574 TCTGAATTTGCAAGCCTTAATGAATACAGAAAGCAAAATATCCAAATGACCGTGAAGGCG 3633
    |||||
DB 1801 TCTGAATTTGCAAGCCTTAATGAATACAGAAAGCAAAATATCCAAATGACCGTGAAGGCG 1860
OY 3634 CCTAGCTCAACAAGTCTT----- 3651
    |||||
DB 1861 CCTAGCTCAACAAGTCTTGAAGTTTGTGATGACACACACACCGCTTACTCTTACG 1920
OY 3652 -----CTTGCTGTGTGACACCCCAATGCGATTAATTTCCGCAATGGTTTTGGA 3702
    |||||
DB 1921 ACAGCAGAGCTTGTGTGTGACACCCCAAGGATTAATTTCCCAATGGTTTTGGA 1980
OY 3703 TTGCTATTGTCAGCTCAATGTTGTATTAATGTAAGAGCTTCTGAGTCTTCTGAAGA 3762
    |||||
DB 1981 TTGCTATTGTCAGCTCAATGTTGTATTAATGTAAGAGCTTCTGAGTCTTCTGAAGA 2040
OY 3763 CGAAGATCTATCAAAATCAAGAAGCTTTGATTTAGATGTTGCTGTAAAGAAATATA 3822
    |||||
DB 2041 CGAAGATCTATCAAAATCAAGAAGCTTTGATTTAGATGTTGCTGTAAAGAAATATA 2100
OY 3823 GATGATCTCAATCATGTGATTTGAATGTGTGACAAAGCTTTTCCGCGCCGGGTAGAGT 3882
    |||||
DB 2101 GATGATCTCAATCATGTGATTTGAATGTGTGACAAAGCTTTTCCGCGCCGGGTAGAGT 2160
OY 3883 GGCAATGCTCTTATGGAATTAACCTTAATGAGGCTTTATGTGCTTCAAGAGCAAT 3942
    |||||
DB 2161 GGCAATGCTCTTATGGAATTAACCTTAATGAGGCTTTATGTGCTTCAAGAGCAAT 2220
OY 3943 TCTCTGAGGAGACAGTGAAGAAGTGAATATGATCATGAAAACTCAACCTCTATTTA 4002
    |||||
DB 2221 TCTCTGAGGAGACAGTGAAGAAGTGAATATGATCATGAAAACTCAACCTCTATTTA 2280
OY 4003 GATTCGTAAATGAAAACCGAGTTTGTGTAATATTCCTGCTGTGAAAACCTTTAAAGT 4062
    |||||
DB 2281 GATTCGTAAATGAAAACCGAGTTTGTGTAATATTCCTGCTGTGAAAACCTTTAAAGT 2340
OY 4063 TCAATATCCCAAGATGCTTCAGTGTCCATAGTGAATTAATGAGCGCAAGAGCAGGCG 4122
    |||||
DB 2341 TCAATATCCCAAGATGCTTCAGTGTCCATAGTGAATTAATGAGCGCAAGAGCAGGCG 2400
OY 4123 GTGAGAAGTTACAACCTGTAAGTGAAGTGTCTCTGTGACCTTTGAGAGTATGCCAG 4182
    |||||
DB 2401 GTGAGAAGTTACAACCTGTAAGTGAAGTGTCTCTGTGACCTTTGAGAGTATGCCAG 2460
OY 4183 GGCTGCGCTCTTGTGAGAGTGAAGTTCAGGCTCCCATCATCACTTTCAGTCAATTTT 4242
    |||||
DB 2461 GGCTGCGCTCTTGTGAGAGTGAAGTTCAGGCTCCCATCATCACTTTCAGTCAATTTT 2520
OY 4243 ATTTTCTGTTTGAAGTTCTGTACTTATGAACTTTGGCTGTGA 4287
    |||||
DB 2521 ATTTTCTGTTTGAAGTTCTGTACTTATGAACTTTGGCTGTGA 2565

```

```

RESULT 13
AK095888
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

REFERENCE
AUTHORS
Oshima,A., Takahashi-Fujii,A., Tanase,T., Imose,N., Takeuchi,K.,
Arita,M., Mutsaers,K., Yuki,H., Hara,H., Sugiyama,T., Irie,R.,
Otsuki,T., Sato,H., Wakamatsu,A., Ishii,S., Yamamoto,J., Isono,Y.,
Kawai-Hio,Y., Saito,K., Nishikawa,T., Kimura,K., Yamashita,H.,
Matsuo,K., Nakamura,Y., Sekine,M., Kikuchi,H., Kanda,K.,
Wagatsuma,M., Murakawa,K., Kanehori,K., Sugiyama,A., Kawakami,B.,
Suzuki,Y., Sugano,S., Nagahara,K., Masuno,Y., Nagai,K. and
Isogai,T.
TITLE
NEO human cDNA sequencing project
JOURNAL
Unpublished
REFERENCE
2 (bases 1 to 2273)
AUTHORS
Isogai,T. and Yamamoto,J.
TITLE
Direct Submission
JOURNAL
Submitted (04-JUN-2002) Takao Isogai, FLJ Project (HRI Team) ; 2-6-7
Kazusa-Kamatari, Kisarazu, Chiba 292-0832, Japan
(E-mail:genomics@hri.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986)
NEO human cDNA sequencing project supported by Ministry of
Economy, Trade and Industry of Japan; cDNA full insert sequencing:
Research Association for Biotechnology (RAB); cDNA library
construction: Helix Research Institute (HRI) (supported by Japan
Key Technology Center etc.); 5'- & 3'-end one pass sequencing: RAB,
HRI, and Biotechnology Center, National Institute of Technology and
Evaluation; clone selection for full insert sequencing: HRI and
RAB; annotation: HRI and RAB.
FEATURES
source
1..2273
location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="HCHON2006459"
/cell_type="chondrocytes (HC)"
/clone_lib="HCHON2"
/notes="cloning vector: pME18FLJ3-primary culture,
chondrocytes"
66..2063
/notes="unnamed protein product"
/codon_start=1
/protein_id="BAC04642.1"
/db_xref="GI:21755237"
/translation="MQGPPPLTAAHLTCVTAALAVADGPFLVTAPGIIIRPGNVTI
GVELLHCPQVTVKAEILLKASNLTVLAEVFEKSGFETLPLPLNSADEIY
ELRVGTRODELFLFNSRLSPETFRKIFVPIOTDKALVPRKOEVRVLTSPDKPY
KTSNILLKDKSNLLOOMLSQOGLGVISKTFQLSHPIDIGMSIYOVDQTYVOS
FOVSEYVLPKREYVTLQTPLYCSMNSKHLNGITTAQTYGKPKGVNLTPLPSWKG
KNKITKTKINGANFSEFNDEMKVMSNGLSFYLDLSSGPVELILTIVESVTGI
SRNVTNVEFKQHDYILIEFPDYTVLAKSLNFTAVKTRADGNDLTLEBRANNVIT
VTORVNTYMGSGNSGNOKMEAVOKINTVQSGTFKLEPFLBSSSELOLVAFLGS
KSSMAVHSLFKSPSKTYIOLKTRDENIKVSGPELVGNGKRLKELSYVNSRGDLVA
VGKONSMTSLTPNSWTPKACVITYIYEDGETISDYLKIPVOLFNRKIKLVMSKV
KAERKESYLSISYTOPDSIYGIYAVDPSVMSNLSNITMENVEHLELVNTGYLLG
MFNNSFANFOEGMLVLTDAINTKDYIDGVINDLFTGDEAL"
BASE COUNT
712 a 431 c 462 g 668 t
ORIGIN
Query Match
Best Local Similarity 99.9%; Pred. No. 0; Length 2273;
Matches 1970; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
OY 1 ATGACGGGCCCAACCGCTCTGACCGCGCCCAACCTCTCTGTGCTGACGCGCGCTG 60
    |||||

```

Db 66 ATGACGAGCCCAACCGCTCTGACCGCCGCACTCTGCGTGTGACCGCGCGCTG 125  
 Qy 61 GCCGCGCTCCCGGCGCTCGGTTTCTGGTGA CAGCCCCAGGAGATCATCAGGCCCGGAGGA 120  
 Db 126 GCCGTGGCTCCCGGCGCTCGGTTTCTGGTGA CAGCCCCAGGAGATCATCAGGCCCGGAGGA 185  
 Qy 121 AATGTGACTATTTGGGGTGGAGCTTCTGTGAACA CTGCCCTTGCACAGTGA CTGTGAAGCG 180  
 Db 186 AATGTGACTATTTGGGGTGGAGCTTCTGTGAACA CTGCCCTTGCACAGTGA CTGTGAAGCG 245  
 Qy 181 GAGCTGCTCAAGACAGCATCAAACTCATCTGTCTGTCTGTCTGTGAAGCAGAAAGAGTCTTT 240  
 Db 246 GAGCTGCTCAAGACAGCATCAAACTCATCTGTCTGTCTGTGAAGCAGAAAGAGTCTTT 305  
 Qy 241 GAAAAAGGCTCTTTTAAGACACTTACTCTTCATCACTACTCTGAACAGAGCAGATGAG 300  
 Db 306 GAAAAAGGCTCTTTTAAGACACTTACTCTTCATCACTACTCTGAACAGAGCAGATGAG 365  
 Qy 301 ATTATGAGCTACGTTGAACCGGACGTACCCAGAGTGAATTTATCTCTAATAGTACC 360  
 Db 366 ATTATGAGCTACGTTGAACCGGACGTACCCAGAGTGAATTTATCTAATAGTACC 425  
 Qy 361 CGCTTATCATTTGAGACCAAGAAATATCTGTCTTCAATCAACAGACAGGCTTATAC 420  
 Db 426 CGCTTATCATTTGAGACCAAGAAATATCTGTCTTCAATCAACAGACAGGCTTATAC 485  
 Qy 421 AAGCAAGCAAGAGTGAAGTTGCACTTGTTACACTCTCTGAGATTTTAAAGCTTAC 480  
 Db 486 AAGCAAGCAAGAGTGAAGTTGCACTTGTTACACTCTCTGAGATTTTAAAGCTTAC 545  
 Qy 481 AAAACCTCTTTAAACATTTCTCAATTAAGAGCCCAATCAATTTGATCCAAAGTGTG 540  
 Db 546 AAAACCTCTTTAAACATTTCTCAATTAAGAGCCCAATCAATTTGATCCAAAGTGTG 605  
 Qy 541 TCACAAACAAGTGAATCTTGAGAGTCATTTCCAAACTTTACGATCTTCCATCCATA 600  
 Db 606 TCACAAACAAGTGAATCTTGAGAGTCATTTCCAAACTTTACGATCTTCCATCCATA 665  
 Qy 601 CTGTGTGCTGTCTATTCAGTTCAGTGAATGACCAAGATTTATCATCATTTACG 660  
 Db 666 CTGTGTGCTGTCTATTCAGTTCAGTGAATGACCAAGATTTATCATCATTTACG 725  
 Qy 661 GTTTCAGAAATGTAATACCAAAATTTGAAGTGACTTTGCACAGACCATTAATTTGTTCT 720  
 Db 726 GTTTCAGAAATGTAATACCAAAATTTGAAGTGACTTTGCACAGACCATTAATTTGTTCT 785  
 Qy 721 ATGAATTCGAACATTTAATGTAATGTAATCAACGCGAAAGTATACATATGGAAGCAGTG 780  
 Db 786 ATGAATTCGAACATTTAATGTAATGTAATCAACGCGAAAGTATACATATGGAAGCAGTG 845  
 Qy 781 AAAAGAGACGTAAAGCTTACATTTTACCTTATCTTTGGGAAAGAAATTAATTT 840  
 Db 846 AAAAGAGACGTAAAGCTTACATTTTACCTTATCTTTGGGAAAGAAATTAATTT 905  
 Qy 841 ACAAAAACATTTAAGATTAATGATCTGCAAACTCTCTTTAATGATGAAGATGAAA 900  
 Db 906 ACAAAAACATTTAAGATTAATGATCTGCAAACTCTCTTTAATGATGAAGATGAAA 965  
 Qy 901 AATGTATGATCTTCAATATGACTTTTGTGAATACCTGATCTATCTTCCCTGAGCA 960  
 Db 966 AATGTATGATCTTCAATATGACTTTTGTGAATACCTGATCTATCTTCCCTGAGCA 1025  
 Qy 961 GTATAAATTTTAACACAGTGAACGAATCACTTACAGTATTTCAAGAAAGTGAAGACT 1020  
 Db 1026 GTATAAATTTTAACACAGTGAACGAATCACTTACAGTATTTCAAGAAAGTGAAGACT 1085  
 Qy 1021 AATGTGTTCTCAAGCAACATGATTAACATCAATGATTTTGTATTAATCACTGTCCTG 1080  
 Db 1086 AATGTGTTCTCAAGCAACATGATTAACATCAATGATTTTGTATTAATCACTGTCCTG 1145  
 Qy 1081 AAGCATCTCTCAACTTCACAGCCACTGTGAAGGTAACCTGTGTGATGGCAACCAACTG 1140  
 Db 1146 AAGCATCTCTCAACTTCACAGCCACTGTGAAGGTAACCTGTGTGATGGCAACCAACTG 1205

Qy 1141 ACTCTTGAAAGAAAGAAATTAATGTATGATCAATACAGTGA CACAGAGAACTAATCTAG 1200  
 Db 1206 ACTCTTGAAAGAAAGAAATTAATGTATGATCAATACAGTGA CACAGAGAACTAATCTAG 1265  
 Qy 1201 TACTGAGCGGATCTTAACACTGGAATACAGAAATGGAAGCTGTTCAGAAATTAATAT 1260  
 Db 1266 TACTGAGCGGATCTTAACACTGGAATTAAGAAATGGAAGCTGTTCAGAAATTAATAT 1325  
 Qy 1261 ACTGTCCCCAAAGTGAACCTTTTAAGATTGAATTTCCATCTCGAGAGATTCAGTGA 1320  
 Db 1326 ACTGTCCCCAAAGTGAACCTTTTAAGATTGAATTTCCATCTCGAGAGATTCAGTGA 1385  
 Qy 1321 CTACAGTTGAAGGCTTAATTTCTTGTGATGAAGTGAAGTGAAGTGAAGTGAAGTGA 1380  
 Db 1386 CTACAGTTGAAGGCTTAATTTCTTGTGATGAAGTGAAGTGAAGTGAAGTGAAGTGA 1445  
 Qy 1381 AAGTCTCTTAATGAACATCAATCACTAATCAATCAATCAATCAATCAATCAATCA 1440  
 Db 1446 AAGTCTCTTAATGAACATCAATCACTAATCAATCAATCAATCAATCAATCAATCA 1505  
 Qy 1441 TCGCTTTTGAAGTGTGTGTTAGTGGCAACAAACGATGAAGAGTTAAGCTAATAGTA 1500  
 Db 1506 TCGCTTTTGAAGTGTGTGTTAGTGGCAACAAACGATGAAGAGTTAAGCTAATAGTA 1565  
 Qy 1501 GTATCCAGGGGACAGTTGTGCTGTAGAAAACAAATCAACAAATGTTCTCTTTAACA 1560  
 Db 1566 GTATCCAGGGGACAGTTGTGCTGTAGAAAACAAATCAACAAATGTTCTCTTTAACA 1625  
 Qy 1561 CCAGAAAATTTCTGAGCTCCAAAACCTGTGTATTTGTATTAATTTGAATGATGG 1620  
 Db 1626 CCAGAAAATTTCTGAGCTCCAAAACCTGTGTATTTGTATTAATTTGAATGATGG 1685  
 Qy 1621 GAAATTAATGATGATGTTCTAATAATTCCTGTGACCTGTTTAAATAATGATGAAG 1680  
 Db 1686 GAAATTAATGATGATGTTCTAATAATTCCTGTGACCTGTTTAAATAATGATGAAG 1745  
 Qy 1681 CTATATTTGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1740  
 Db 1746 CTATATTTGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1805  
 Qy 1741 ACAAGCTGACTCCATGATGTTGGAGTTGAGTGTGACAAAGTGAATCTGATGAAT 1800  
 Db 1806 ACAAGCTGACTCCATGATGTTGGAGTTGAGTGTGACAAAGTGAATCTGATGAAT 1865  
 Qy 1801 GCCTCTTAATGATTTCAATGGAATATGAGTGTGACAAAGTGAAGTGAAGTGAAGTGA 1860  
 Db 1866 GCCTCTTAATGATTTCAATGGAATATGAGTGTGACAAAGTGAAGTGAAGTGAAGTGA 1925  
 Qy 1861 TATTAATTTAGGATGATTCATGAATTTCTTTGACAGTCTTTGACAGAAATGAGTGA 1920  
 Db 1926 TATTAATTTAGGATGATTCATGAATTTCTTTGACAGTCTTTGACAGAAATGAGTGA 1985  
 Qy 1921 TTGACAGATCAAACTCAGAGAGATTAATGATGATGATGATGATGATGATGAT 1971  
 Db 1986 TTGACAGATCAAACTCAGAGAGATTAATGATGATGATGATGATGATGATGAT 2036

RESULT 14  
 AX083146  
 LOCUS 1300 bp DNA linear PAT 28-FEB-2001  
 DEFINITION Sequence 49 from Patent WO0110903.  
 AX083146  
 VERSION  
 AX083146.1 GI:13185055  
 KEYWORDS  
 SOURCE  
 ORGANISM  
 Homo sapiens (human)  
 Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 REFERENCE  
 AUTHORS Yue,H., Lai,P., Tang,Y.T., Bandman,O., Baughn,M.R., Azimzai,Y.,  
 Lu,D.A. and Yang,J.  
 TITLE Proteases and protease inhibitors

JOURNAL Patent: WO 0110903-A 49 15-FEB-2001;  
 Incyte Genomics, Inc. (US)  
 Location/Qualifiers  
 FEATURES  
 source 1..1300  
 /organism="Homo sapiens"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9606"  
 /note="Incyte ID No: 3658034CB1"  
 BASE COUNT 353 a 239 c 289 g 419 t  
 ORIGIN  
 Query Match 17.0%; Score 730.8; DB 6; Length 1300;  
 Best Local Similarity 93.6%; Pred. No. 1.3e-148;  
 Matches 805; Conservative 0; Mismatches 2; Indels 53; Gaps 2;  
 Db 3479 AGGGAATCCCAATATATGAGTGGCTTAAGAGGCAAGAAATATGCTGGGAGTTTGCAT 3538  
 1 AGGGAATCCCAATATATGAGGCGGCTTAAGAGGCAAGAAATATGCTGGGAGTTTGCAT 60  
 Oy 3539 CTACTCAGAGATACCACTGTGGCTTTAAAGGCTCTGTCTGAATTTTGAGCCCTAATGATA 3598  
 Db 61 CTACTCAGAGATACCACTGTGGCTTTAAAGG--CTGTCTGAATTTTGAGCCCTAATGATA 118  
 Oy 3599 CAGAAAGCAAAATATCCAAAGTGAACGCTGAACGGGCTTACCTCAGCAAGTCTT----- 3651  
 Db 119 CAGAAAGCAAAATATCCAAAGTGAACGCTGAACGGGCTTACCTCAGCAAGTCTTAAAGT 178  
 Oy 3652 -----CTTGCTGTGTACAGC 3667  
 Db 179 TTCTGATTAACACACACACACCGCTTACCTCTTACAGACAGAGCTTGTGTGTACAGC 238  
 Oy 3668 CAATGCGCATTAATATTTCCGCAATATGTTTGATTTGCTATTTGTCAGCTCAATGTTG 3727  
 Db 239 CAATGCGCATTAATATTTCCGCAATATGTTTGATTTGCTATTTGTCAGCTCAATGTTG 298  
 Oy 3728 TATATATGTGAAGGCTTCTGGCTTTCTAGAAAGCAAGATCTATCCAAATCAAGAG 3787  
 Db 299 TATATATGTGAAGGCTTCTGGCTTTCTAGAAAGCAAGATCTATCCAAATCAAGAG 358  
 Oy 3788 CCTTTGATTTAGATGTGTGTAAAGAAATTAAGATATCTCATATCTATGATTTGA 3847  
 Db 359 CCTTTGATTTAGATGTGTGTAAAGAAATTAAGATATCTCATATCTATGATTTGA 418  
 Oy 3848 ATGTGTGAACAAGCTTTCCGGCCCGGTAGAGTGGCATGGCTTTATGAAAGTTAAC 3907  
 Db 419 ATGTGTGAACAAGCTTTCCGGCCCGGTAGAGTGGCATGGCTTTATGAAAGTTAAC 478  
 Oy 3908 TATTAAGTGGCTTTATGTGTCTTCAAGACAAATTTCTGTGACGAGACAGTGAAGAG 3967  
 Db 479 TATTAAGTGGCTTTATGTGTCTTCAAGACAAATTTCTGTGACGAGACAGTGAAGAG 538  
 Oy 3968 TGAATATATATATGAGAAATCAACCTTATTTAGATTTCTGAATTAAGAAACCCAGTTT 4027  
 Db 539 TGAATATATATATGAGAAATCAACCTTATTTAGATTTCTGAATTAAGAAACCCAGTTT 598  
 Oy 4028 GTGTTAATATCTCTGTGTGAGAAATTTAAAGTTCAATATCCCAAGATGCTCAGTGT 4087  
 Db 599 GTGTTAATATCTCTGTGTGAGAAATTTAAAGTTCAATATCCCAAGATGCTCAGTGT 658  
 Oy 4088 CCATATGATTAATCTATGAGCAAGAGAGACAGCGGTGAGAAATTAACAATCTGAAATGA 4147  
 Db 659 CCATATGATTAATCTATGAGCAAGAGAGACAGCGGTGAGAAATTAACAATCTGAAATGA 718  
 Oy 4148 AGCTGTCTCTGTGTGACCTTTGAGAGTGTCAAGGCTCCGCTCTTGTGTGAGATGAG 4207  
 Db 719 AGCTGTCTCTGTGTGACCTTTGAGAGTGTCAAGGCTCCGCTCTTGTGTGAGATGAG 778  
 Oy 4208 CTTCAGGCTCCATCATCACTTTGATGATTTTATTTTCTGTTCAAGCTTCTGATCT 4267  
 Db 779 CTTCAGGCTCCATCATCACTTTGATGATTTTATTTTCTGTTCAAGCTTCTGATCT 838  
 Oy 4268 TTATGGAACCTTGGCTGTGA 4287  
 |||||

Db 839 TTATGGAACCTTGGCTGTGA 858  
 RESULT 15  
 AL590428  
 LOCUS  
 DEFINITION  
 Human DNA sequence from clone RP11-553A21 on Chromosome 6, complete  
 sequence.  
 ACCESSION  
 AL590428 AC026605  
 VERSION  
 AL590428.7 GI:15072593  
 KEYWORDS  
 HTG.  
 SOURCE  
 Homo sapiens (human)  
 ORGANISM  
 Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 1 (bases 1 to 163577)  
 REFERENCE  
 AUTHORS  
 TITLE  
 JOURNAL  
 COMMENT  
 During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.  
 This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred-quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one N13 subclone; and the assembly was confirmed by restriction digest. The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em., EMBL; Sw., SWISSPROT; Tr., TrEMBL; Wp., WORMPEP; Information on the WORMPEP database can be found at  
 http://www.sanger.ac.uk/Projects/C\_elegans/wormpep This sequence was generated from part of bacterial clone contigs of human chromosome 6, constructed by the Sanger Centre Chromosome 6 Mapping Group. Further information can be found at  
 http://www.sanger.ac.uk/HGP/Chr6  
 RP11-553A21 is from the library RPCT-11.2 constructed by the group of Pieper de Jong. For further details see  
 http://www.chori.org/bacpac/home.htm  
 VECTOR: pBAC3.6  
 IMPORTANT: This sequence is not the entire insert of clone RP11-553A21. It may be shorter because we sequence overlapping sections only once, except for a 100 base overlap.  
 The true right end of clone RP11-553A21 is at 163577 in this sequence. The true left end of clone RP11-525G3 is at 88067 in this sequence. The true right end of clone RP3-197H23 is at 2000 in this sequence.  
 FEATURES  
 source  
 location/Qualifiers  
 1..163577  
 /organism="Homo sapiens"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9606"  
 /chromosome="6"  
 /clone="RP11-553A21"  
 /clone\_1ib="RPCT-11.2"  
 285..433  
 /note="AluNb repeat: matches 1..119 of consensus"  
 434..743  
 /note="AluY repeat: matches 1..311 of consensus"  
 744..919  
 /note="AluY repeat: matches 1..305 of consensus"  
 988..1075  
 /note="L2 repeat: matches 2626..2708 of consensus"  
 1166..1384  
 /note="L1MC5 repeat: matches 7317..7536 of consensus"  
 repeat\_region  
 repeat\_region  
 repeat\_region  
 repeat\_region  
 repeat\_region

repeat_region	1385..1690	/note="AluXc repeat: matches 1..306 of consensus"
repeat_region	1691..1759	/note="AluXc repeat: matches 7536..7604 of consensus"
repeat_region	1760..2024	/note="AluY repeat: matches 35..287 of consensus"
repeat_region	2025..2248	/note="L1MC5 repeat: matches 7604..7887 of consensus"
repeat_region	2588..2706	/note="L1MC/D repeat: matches 5389..5514 of consensus"
misc_feature	3686..4489	/note="Cpg island"
repeat_region	3695..3736	/evidence=not_experimental
repeat_region	4691..4996	/note="21 copies 2 mer cc 78% conserved"
repeat_region	5628..5977	/note="AluIo repeat: matches 3..306 of consensus"
repeat_region	5978..6030	/note="MLT7FB repeat: matches 1..403 of consensus"
repeat_region	6118..6474	/note="MER3A repeat: matches 44..37 of consensus"
repeat_region	6475..6771	/note="L1MC3 repeat: matches 7398..7739 of consensus"
repeat_region	6772..7373	/note="AluY repeat: matches 1..290 of consensus"
repeat_region	7491..7764	/note="L1MC3 repeat: matches 6740..7398 of consensus"
repeat_region	7887..8166	/note="AluIo repeat: matches 1..286 of consensus"
repeat_region	8365..8430	/note="AluXc repeat: matches 1..287 of consensus"
repeat_region	8431..8729	/note="33 copies 2 mer tc 89% conserved"
repeat_region	8779..8975	/note="AluSq repeat: matches 1..302 of consensus"
repeat_region	8976..9292	/note="MER3J repeat: matches 129..323 of consensus"
repeat_region	9293..9415	/note="AluXc repeat: matches 1..311 of consensus"
repeat_region	9516..9647	/note="MER3J repeat: matches 1..129 of consensus"
repeat_region	9698..9994	/note="MIR repeat: matches 106..226 of consensus"
repeat_region	10005..10048	/note="AluXc repeat: matches 1..295 of consensus"
repeat_region	10050..10999	/note="22 copies 2 mer tc 79% conserved"
misc_feature	10622..11798	/note="SVA repeat: matches 3..954 of consensus"
repeat_region	11056..11951	/note="Cpg island"
repeat_region	12038..12152	/evidence=not_experimental
repeat_region	12561..12844	/note="SVA repeat: matches 519..1386 of consensus"
repeat_region	12928..13220	/note="L2 repeat: matches 2588..2707 of consensus"
repeat_region	13294..13417	/note="AluXc repeat: matches 3..287 of consensus"
repeat_region	13451..13636	/note="AluXc repeat: matches 2..295 of consensus"
repeat_region	13637..13697	/note="FLM3 C repeat: matches 1..123 of consensus"
repeat_region	13926..14215	/note="AluSq repeat: matches 1..189 of consensus"
repeat_region	13926..14215	/note="AluY repeat: matches 1..63 of consensus"
repeat_region	14216..14662	/note="MER39 repeat: matches 15..148 of consensus"
repeat_region	14688..14735	/note="AluY repeat: matches 1..288 of consensus"
repeat_region	14688..14735	/note="MER39 repeat: matches 148..543 of consensus"

	repeat_region	/note="LTR29 repeat: matches 571. .619 of consensus" 14332. .15863 /note="LIM4 repeat: matches 5262. .6295 of consensus" 15978. .16180 /note="L2 repeat: matches 2208. .2414 of consensus" 16387. .16511 /note="L2 repeat: matches 2656. .2734 of consensus" 17252. .17283 /note="L16 copies 2 mer ac 96% conserved" 17372. .17674 /note="AluSg repeat: matches 1. .303 of consensus" 18061. .19334 /note="LIM10 repeat: matches 5026. .6320 of consensus" 19352. .19437 /note="LIM3C repeat: matches 734. .820 of consensus" 19545. .19620 /note="LIM3C repeat: matches 922. .997 of consensus" 20319. .20442 /note="AluIo/FRM repeat: matches 179. .305 of consensus" 20921. .21277 /note="THE1B repeat: matches 1. .364 of consensus" 21480. .21775 /note="AluSg repeat: matches 1. .286 of consensus" 21799. .21916 /note="FLM1C repeat: matches 1. .118 of consensus" 22546. .22713 /note="AluSg/x repeat: matches 87. .262 of consensus" 22895. .23001 /note="LIM2 repeat: matches 5827. .5924 of consensus" 23005. .23303 /note="AluSx repeat: matches 14. .312 of consensus" 23429. .23525 /note="AluSg/x repeat: matches 213. .309 of consensus" 23528. .23628 /note="AluSg/x repeat: matches 12. .112 of consensus" 23626. .23695 /note="AluSx repeat: matches 62. .142 of consensus" 23696. .23996 /note="AluSg repeat: matches 1. .304 of consensus" 23997. .24222 /note="AluSx repeat: matches 142. .309 of consensus" 24443. .24416 /note="AluSg/x repeat: matches 127. .300 of consensus" 24427. .24734 /note="AluSx repeat: matches 1. .300 of consensus" 24807. .25158 /note="MLT1A1 repeat: matches 1. .363 of consensus" 25179. .25484 /note="AluSx repeat: matches 1. .299 of consensus" 25494. .25519 /note="L13 copies 2 mer at 100% conserved" 25614. .25730 /note="FLM1C repeat: matches 2. .118 of consensus" 26558. .27263 /note="AluIo repeat: matches 7. .311 of consensus" 27497. .27814 /note="LIM4 repeat: matches 4499. .4833 of consensus" 27866. .27943 /note="MLT1U repeat: matches 410. .492 of consensus"		
Query Match	5.4%;	Score 232;	DB 9;	Length 163577;
Best Local Similarity	100.0%;	Pred. No. 6.9e-40;		
Matches 232;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
337	CTACCTTGAAACAGTGCAGATGAGATTTATAGAGCTACGCTGTAACCGGAGCTACCCAGAT	336		
337	GAGATTTATTCCTATATAGTACCGGCTATATCATTTGAGACCAAGAGATATCTGCTTC	396		
Db	80610 GAGATTTATTCCTATATAGTACCGGCTATATCATTTGAGACCAAGAGATATCTGCTTC	806659		
397	ATTCAACAGACAGCGCTTATACAGCCAAAGCAGAGAGTAGATTTTCGATTTGTACA	456		

Db 80670 ATTCAACAGACAGGCGCTTATACAGCCAAAGCAAGAGTGAAGTTCCGATTGTTACA 80729  
Qy 457 CTCTCTCAGATTTTAAGCCTTACAAACCTTTTAAACATTCTCATTAAG 508  
Db 80730 CTCTCTCAGATTTTAAGCCTTACAAACCTTTTAAACATTCTCATTAAG 80781

Search completed: January 16, 2004, 18:48:41  
Job time : 15013 secs

**THIS PAGE BLANK (USPTO)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 16, 2004, 12:36:15 ; Search time 1017 Seconds  
(without alignments)  
11379.047 Million cell updates/sec

Title: US-10-020-095-3

Perfect score: 4287  
Sequence: 1 atcgagggccacgcgcctcct.....ttatggaacttgcctgtga 4287

Scoring table: IDENTITY NUC  
Gapop 10.0, Gapext 1.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Database :

N\_Geneseq\_19Jun03.\*  
1: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT.\*  
2: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT.\*  
3: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT.\*  
4: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT.\*  
5: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT.\*  
6: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT.\*  
7: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT.\*  
8: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT.\*  
9: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT.\*  
10: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT.\*  
11: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT.\*  
12: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT.\*  
13: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT.\*  
14: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT.\*  
15: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT.\*  
16: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT.\*  
17: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT.\*  
18: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT.\*  
19: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.\*  
20: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.\*  
21: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.\*  
22: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.\*  
23: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.\*  
24: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT.\*  
25: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2003.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	4281.2	99.9	4369	25	AAD49434 Human r150 gene #1
2	4224.4	98.5	4761	24	ABQ79964 Human CD109 K1 pro
3	4224.4	98.5	5882	25	AAD49440 Human blood cell s
4	4224.4	98.5	5895	24	ABQ79965 Human CD109 K1-H7
5	4223.4	98.5	5883	25	AAD49435 Human r150 DNA #1.
6	4222.8	98.5	4761	24	ABQ79965 Human CD109 K1 var
7	4222.8	98.5	5895	24	ABQ79967 Human CD109 K1-H7
8	4221.4	98.5	4335	24	AAL49815 Human platelet all

9	4219.8	98.4	4335	24	AAL49816
10	4219.6	98.4	4473	21	AAA62010
11	4216.6	98.4	4435	21	AAA60199
12	4140.2	96.6	4146	25	AAD49436
13	4082.4	95.2	4197	25	AAD49437
14	3187.4	74.4	3535	25	ABQ79968
15	3185.8	74.3	3535	24	ABQ79969
16	1017	23.7	1448	24	ABL90478
17	975.4	22.8	1459	22	ABA08825
18	812.2	18.9	2403	24	ABV77330
19	730.8	17.0	1300	22	AAF81735
20	541.6	12.6	875	20	AAX20449
21	525	12.2	821	25	ABX08843
22	497.6	11.6	744	20	AAZ15521
23	331.6	7.7	690	23	AA570864
24	300	7.0	300	20	AAZ15539
25	272	6.3	354	20	ABX41702
26	175	4.1	744	25	ABX41702
27	174	4.1	4527	22	AAC90951
28	174	4.1	4527	22	AAC90952
29	172.6	4.0	2285	23	ABL20611
30	171.6	4.0	2480	23	ABL2153
31	152.2	3.6	3612	23	ABL19705
32	143.8	3.4	6324	23	ABL19704
33	127.2	3.0	5611	23	ABL20610
34	127.2	3.0	6979	23	ABL2152
35	126.8	3.0	4615	23	AA592193
36	126.8	3.0	4615	24	ABN97232
37	126.8	3.0	4629	23	AA592192
38	126.8	3.0	5816	21	AACT4484
39	125.2	2.9	4487	24	ABK92035
40	125.2	2.9	4488	24	AB559327
41	120.8	2.8	4660	24	ABO93897
42	120.6	2.8	2608	24	AAL49817
43	119.8	2.8	4501	25	ABX70464
44	119.4	2.8	1140	22	ABA06696
45	119.4	2.8	1140	22	ABA06869

#### ALIGNMENTS

RESULT 1  
AAD49434  
ID--AAD49434 standard; DNA; 4369 BP.

AC	AAD49434;	
DT	24-MAR-2003 (first entry)	
XX		
DE	Human r150 gene #1.	
XX		
KW	Human; GPI-anchored TGF-beta1 binding protein; r150 protein; cancer;	
KW	glycosylphosphatidylinositol; transforming growth factor-beta1;	
XX	therapy; gene; de.	
OS	Homo sapiens.	
XX		
FH	Key	Location/Qualifiers
FT	CDS	26..4312
FT		/tag= a
FT		/product= "Human r150 protein #1"
FT		/transl_except= (pos:2132..2134, aa:Xaa)
FT		/note= "Xaa corresponds to Ser, Tyr"
XX		
PN	WO200285942-A2.	
XX		
PD	31-OCT-2002.	
XX		
PF	24-APR-2002; 2002WO-CA00560.	
XX		
PR	24-APR-2001; 2001US-285713P.	
PR	14-FEB-2002; 2002US-356163P.	

Human platelet all  
Hydrophobic domain  
Hydrophobic domain  
Human blood cell s  
Human r150 DNA #2.  
Human CD109 K15 pr  
Human CD109 K15 va  
Human polynucleoti  
Human secreted pro  
Alpha-1 proteinase  
Human protease and  
Human secreted pro  
Angiogenesis-asoc  
Human gene express  
DNA encoding novel  
Human gene express  
Bovine EST associa  
Human gene express  
C. elegans alpha-2  
Drosophila melanog  
Drosophila melanog  
Drosophila melanog  
Drosophila melanog  
Drosophila melanog  
Drosophila melanog  
Gene #3730 used to  
DNA encoding novel  
Human ORF ORF39 P  
DNA encoding novel  
Human alpha-2-macr  
Human pregnancy zo  
Human platelet all  
DNA encoding human  
Human cDNA SEQ ID  
Human polynucleoti

XX (UYMC-) UNIV MCGILL.  
 PA Philip A, Tam B;  
 PI WPI: 2003-093100/08.  
 DR P-PSDB; AAE32012.  
 XX Novel transforming growth factor (TGF)-beta 1 binding reagent which  
 PT comprises r150 protein which acts as accessory receptor of TGF-beta,  
 PT useful for negatively modulating TGF-beta activity, and thus for  
 PT treating cancer -  
 XX Claim 7; Page 91-93; 127pp; English.  
 PS The invention relates to novel transforming growth factor (TGF)-beta1  
 CC binding reagent which comprises a glycosylphosphatidylinositol (GPI)-  
 CC anchored TGF-beta1 binding protein referred to as r150 which acts as  
 CC an accessory receptor of TGF-beta. The invention is used for negatively  
 CC modulating TGF-beta activity, and thus for treating conditions  
 CC characterised by overproduction of TGF-beta, such as cancer. Antisense  
 CC molecules of the invention are used for increasing TGF-beta availability  
 CC and increase graft success. The present sequence is a gene encoding  
 CC human r150 protein.  
 CC  
 SQ Sequence 4369 BP; 1301 A; 842 C; 932 G; 1293 T; 1 other;  
 Query Match 99.9%; Score 4281.2; DB 25; Length 4369;  
 Basic Local Similarity 99.9%; Pred. No. 0;  
 Matches 4283; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Db 626 CTGTGACCTGCTATTCAGATTCAAGTGAATGACACAGACATCATTCATTCAG 685  
 Qy 661 GTTTCAGAAATGATGATTAACCAAAATTTGAAGTCTTGGACACCATTAATGTTCT 720  
 Db 686 GTTTCAGAAATGATGATTAACCAAAATTTGAAGTCTTGGACACCATTAATGTTCT 745  
 Qy 721 ATGAATTCAGAAATGATGATTAACCAAAATTTGAAGTCTTGGACACCATTAATGTTCT 780  
 Db 746 ATGAATTCAGAAATGATGATTAACCAAAATTTGAAGTCTTGGACACCATTAATGTTCT 805  
 Qy 781 AAGGAGAGCTAACCGCTTACATTTTACCTTTTCTTTGGGAAAGAAATATTT 840  
 Db 806 AAGGAGAGCTAACCGCTTACATTTTACCTTTTCTTTGGGAAAGAAATATTT 865  
 Qy 841 ACAAACATTTAAGTAAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 900  
 Db 866 ACAAACATTTAAGTAAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 925  
 Qy 901 AATGTAATGATTTCTCAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 960  
 Db 926 AATGTAATGATTTCTCAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 985  
 Qy 961 GTAGAAATTTTAAACCAAGTACAGATGATGATGATGATGATGATGATGATGATGATGAT 1020  
 Db 986 GTAGAAATTTTAAACCAAGTACAGATGATGATGATGATGATGATGATGATGATGATGAT 1045  
 Qy 1021 AATGTTCTTCAAGTAAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1080  
 Db 1046 AATGTTCTTCAAGTAAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1105  
 Qy 1081 AAGCATTCTTCAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1140  
 Db 1106 AAGCATTCTTCAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1165  
 Qy 1141 ACTTTGAAGAAAGAAATTAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1200  
 Db 1166 ACTTTGAAGAAAGAAATTAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1225  
 Qy 1201 TACTGAGAGGATCTTAAAGTGAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1260  
 Db 1226 TACTGAGAGGATCTTAAAGTGAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1285  
 Qy 1261 ACTGTCCTTCAAGTAAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1320  
 Db 1286 ACTGTCCTTCAAGTAAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1345  
 Qy 1321 CTACAGTTGAAGGCTTATTTCTTGGTGAATGAATGATGATGATGATGATGATGATGATGATGAT 1380  
 Db 1346 CTACAGTTGAAGGCTTATTTCTTGGTGAATGAATGATGATGATGATGATGATGATGATGATGAT 1405  
 Qy 1381 AAGTCTCTAGTAAAGTAAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1440  
 Db 1406 AAGTCTCTAGTAAAGTAAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1465  
 Qy 1441 TCGCTTTTGAAGTGAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1500  
 Db 1466 TCGCTTTTGAAGTGAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1525  
 Qy 1501 GTATCCAGAGGAGTGAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1560  
 Db 1526 GTATCCAGAGGAGTGAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1585  
 Qy 1561 CCAAGAAATTTCTTGAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1620  
 Db 1586 CCAAGAAATTTCTTGAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1645  
 Qy 1621 GAAATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1680  
 Db 1646 GAAATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1705  
 Qy 1681 CTATATTTGAAGTAAAGTGAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1740  
 Db 1706 CTATATTTGAAGTAAAGTGAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1765

QY 1741 ACACAGCCTGACTCCATAGTTGGGATTTAGCTGTGACAAAAGTGTGATCTGATGAT 1800  
DB 1766 AACAGACCTGACTCCATAGTTGGGATTTAGCTGTGACAAAAGTGTGATCTGATGAT 1825  
QY 1801 GCTCTAATGATATTACATGGAATAATGTGTCCATGAGTTGGAACCTTTATACACAGGA 1860  
DB 1826 GCTCTAATGATATTACATGGAATAATGTGTCCATGAGTTGGAACCTTTATACACAGGA 1885  
QY 1861 TATTTATTTAGGAGATGTTGATGATGATTTGAGAGTTTCCAGAAATGTGACCTGGGTA 1920  
DB 1886 TATTTATTTAGGAGATGTTGATGATGATTTGAGAGTTTCCAGAAATGTGACCTGGGTA 1945  
QY 1921 TTGACAGATGCAAACTCCAGAGATTTATTTGATGATGATTTGATGATGATGATGAT 1980  
DB 1946 TTGACAGATGCAAACTCCAGAGATTTATTTGATGATGATTTGATGATGATGATGAT 2005  
QY 1981 GCTGAGAGGTTTATGAGAGAAATGAGAGCATATTGATGATTTGATGATGATTTCTTTG 2040  
DB 2006 GCTGAGAGGTTTATGAGAGAAATGAGAGCATATTGATGATTTGATGATGATTTCTTTG 2065  
QY 2041 GGTAGAGATCCACATGTCGAAAGCATTTTCCAGAGACTTGATTTGGCTAGACACCAAC 2100  
DB 2066 GGTAGAGATCCACATGTCGAAAGCATTTTCCAGAGACTTGATTTGGCTAGACACCAAC 2125  
QY 2101 ATGGGTTACAGGATTTTCCAGAAATTTGAGATGATGATGATGATGATGATGATGAT 2160  
DB 2126 ATGGGTTACAGGATTTTCCAGAAATTTGAGATGATGATGATGATGATGATGATGAT 2185  
QY 2161 GTGGCTACTGTGTTTGTGATCTCTGAGAGCTGGGCTTTGAGACTTACCACTCCAGTG 2220  
DB 2186 GTGGCTACTGTGTTTGTGATCTCTGAGAGCTGGGCTTTGAGACTTACCACTCCAGTG 2245  
QY 2221 GAGCTCCAGCCTTCCACAACTTTTCAATTTTGTGAACTTCCCTGACCTGTATACGA 2280  
DB 2246 GAGCTCCAGCCTTCCACAACTTTTCAATTTTGTGAACTTCCCTGACCTGTATACGA 2305  
QY 2281 GGTGAGAAATTTGCTTTGAAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2340  
DB 2306 GGTGAGAAATTTGCTTTGAAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2365  
QY 2341 AAGGTATCATTTGAGAAAGTGAACAATTTGATTTCTAATGACTTCAAGTGAATTAAT 2400  
DB 2366 AAGGTATCATTTGAGAAAGTGAACAATTTGATTTCTAATGACTTCAAGTGAATTAAT 2425  
QY 2401 GGCACAGGACCAAGAGACCTTCTGTGTTCCAGAGAGATGGGCAACTGTTCTTTT 2460  
DB 2426 GGCACAGGACCAAGAGACCTTCTGTGTTCCAGAGAGATGGGCAACTGTTCTTTT 2485  
QY 2461 CCCATCAGGCAACACATCTGGGAGAAATTCCTATCAGATCAGCTCTTCAACCACT 2520  
DB 2486 CCCATCAGGCAACACATCTGGGAGAAATTCCTATCAGATCAGCTCTTCAACCACT 2545  
QY 2521 GCTTCTGATGCTGTCAACCAAGATTTTATGAAAGCTGAAAGAAATGAAATCAAT 2580  
DB 2546 GCTTCTGATGCTGTCAACCAAGATTTTATGAAAGCTGAAAGAAATGAAATCAAT 2605  
QY 2581 TCACATTCATCTTATTAAGCTTGAACATGATGATGATGATGATGATGATGATGAT 2640  
DB 2606 TCACATTCATCTTATTAAGCTTGAACATGATGATGATGATGATGATGATGATGAT 2665  
QY 2641 AGTTTCTATTTCTCTCTAATACAGTACGTCGAGTGAAGAAAGTTCAGATCCTGCAAT 2700  
DB 2666 AGTTTCTATTTCTCTCTAATACAGTACGTCGAGTGAAGAAAGTTCAGATCCTGCAAT 2725  
QY 2701 GAGAGATGTTCTTGCTCTTCCATCAATGCTTACCTCATGATGATGATGATGATGATGAT 2760  
DB 2726 GAGAGATGTTCTTGCTCTTCCATCAATGCTTACCTCATGATGATGATGATGATGATGAT 2785  
QY 2761 TGTGTGAAACAGAACTGATTAATTTTGTCTCCAAATTTTACATTTTGTGATTTATCTGACT 2820  
DB 2786 TGTGTGAAACAGAACTGATTAATTTTGTCTCCAAATTTTACATTTTGTGATTTATCTGACT 2845

QY 2821 AAAAAAGCAACGACGATTAATTTGAAAGAAAAGCTTTTCATTTATGAGGCAAGT 2880  
DB 2846 AAAAAAGCAACGACGATTAATTTGAAAGAAAAGCTTTTCATTTATGAGGCAAGT 2905  
QY 2881 TACAGAGAAACTTCTCTATCAGAGGAAATGCTCTTTCAGTCTTTTGGAAATTAAT 2940  
DB 2906 TACAGAGAAACTTCTCTATCAGAGGAAATGCTCTTTCAGTCTTTTGGAAATTAAT 2965  
QY 2941 GACCCCTTGGGAGACATGTTGTGACGCTTTTGTGTTTGAAGTGTTCCTTGAAGCCGAT 3000  
DB 2966 GACCCCTTGGGAGACATGTTGTGACGCTTTTGTGTTTGAAGTGTTCCTTGAAGCCGAT 3025  
QY 3001 CCTTACATGATTAATGATCAGAAATGTGTTACAGAAATACACTTGTGTTAAAGACAT 3060  
DB 3026 CCTTACATGATTAATGATCAGAAATGTGTTACAGAAATACACTTGTGTTAAAGACAT 3085  
QY 3061 CAGAAATCCACGCTGAATTTTGGATTCAGAGAGATGATTTCAATGAGCTTCAAGT 3120  
DB 3086 CAGAAATCCACGCTGAATTTTGGATTCAGAGAGATGATTTCAATGAGCTTCAAGT 3145  
QY 3121 GGCATTAAGTCCAGTACACTTACAGCCTATATGTTAACTTCTCTCGGATATAGA 3180  
DB 3146 GGCATTAAGTCCAGTACACTTACAGCCTATATGTTAACTTCTCTCGGATATAGA 3205  
QY 3181 AAGTATCAGCCTTAATGATGTCAGAGATCTATCCATTTTGGAGTGAATTCAGT 3240  
DB 3206 AAGTATCAGCCTTAATGATGTCAGAGATCTATCCATTTTGGAGTGAATTCAGT 3265  
QY 3241 AGAGGAATTTACAGCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3300  
DB 3266 AGAGGAATTTACAGCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3325  
QY 3301 AGTCTTAAGGAGAAAGCTTTGAAATGTCGACTTGGAGAGCAAGAAAGGTGGC 3360  
DB 3326 AGTCTTAAGGAGAAAGCTTTGAAATGTCGACTTGGAGAGCAAGAAAGGTGGC 3385  
QY 3361 ATGCAATTTCTGGGTGTATCAGAGATCCAACTTTCTGACTCTGGCAGCCAGCTCCCTG 3420  
DB 3386 ATGCAATTTCTGGGTGTATCAGAGATCCAACTTTCTGACTCTGGCAGCCAGCTCCCTG 3445  
QY 3421 GATATTTAGTTGACGCTATGACCTGTCTTCAACTTTTCAATTTTCAACTTTCTGAG 3480  
DB 3446 GATATTTAGTTGACGCTATGACCTGTCTTCAACTTTTCAACTTTTCAACTTTCTGAG 3505  
QY 3481 GGAATCCCAATTTAGAGGTGCTTAAAGGAGAAATTAAGGAGGAGGAGGAGGAGGAGGAG 3540  
DB 3506 GGAATCCCAATTTAGAGGTGCTTAAAGGAGAAATTAAGGAGGAGGAGGAGGAGGAGGAG 3565  
QY 3541 ACTCAGATTAACATGCTGTGTTAAAGGCTGTCTGTAATTTGAGAGCCCTAATGATACGA 3600  
DB 3566 ACTCAGATTAACATGCTGTGTTAAAGGCTGTCTGTAATTTGAGAGCCCTAATGATACGA 3625  
QY 3601 GAAAGCAAAATATCAAGTACCGGTGACGGGCTTACCTCAACAACTCTCTGCTGTG 3660  
DB 3626 GAAAGCAAAATATCAAGTACCGGTGACGGGCTTACCTCAACAACTCTCTGCTGTG 3685  
QY 3661 GTACAGCCCAATGAGGATTAATTTCCGCAATAGTTTGTGATTTGCTATTTGACGCTC 3720  
DB 3686 GTACAGCCCAATGAGGATTAATTTCCGCAATAGTTTGTGATTTGCTATTTGCTATTTGACGCTC 3745  
QY 3721 AATGTTATATTAATGTAAGGCTTCTGGGCTTCTTGAAGAAACGAAGATCTATCCAAAT 3780  
DB 3746 AATGTTATATTAATGTAAGGCTTCTGGGCTTCTTGAAGAAACGAAGATCTATCCAAAT 3805  
QY 3781 CAGAGAGCTTTGATTTAGATGTTGCTGTTAAAGAAATTAAGATGATCTCAATCATGTG 3840  
DB 3806 CAGAGAGCTTTGATTTAGATGTTGCTGTTAAAGAAATTAAGATGATCTCAATCATGTG 3865  
QY 3841 GATTTGATGTTGTGACAGCTTTTCCGCCCCGGGTAGAGATGCTATGATGATGATGATGAT 3900  
DB 3866 GATTTGATGTTGTGACAGCTTTTCCGCCCCGGGTAGAGATGCTATGATGATGATGATGAT 3925  
QY 3901 GTTAACTATTAAGGCTTTATGTTGCTTCAAGAGCAATTTCTCTGAGGAGACAGTG 3960

```

Db      3926 GTTAACTTAAAGTGGCTTTATGGTCTCAGAGCAATTTCTGAGGAGACAGTGTG
Oy      3961 AAGAAAGTAAATATGATCATGAAAACTCAACCTTATAGATTCGTAATGAACC
Db      3986 AAGAAAGTAAATATGATCATGAAAACTCAACCTTATAGATTCGTAATGAACC
Oy      4021 CAGTTTGTGTTAATATCTGCTGTGAGAACTTAAAGTTTCAATACCAAGATGCT
Db      4046 CAGTTTGTGTTAATATCTGCTGTGAGAACTTAAAGTTTCAATACCAAGATGCT
Oy      4081 TCAGTGTCAATAGTGATTAATCAATGAGCAAGGCGGTGAGAGTAACTT
Db      4106 TCAGTGTCAATAGTGATTAATCAATGAGCAAGGCGGTGAGAGTAACTT
Oy      4141 GAAGTAAAGTGTCTCTCTGAGACCTTGGATGATGTCAGAGGCTGCCCTTGTGAG
Db      4166 GAAGTAAAGTGTCTCTCTGAGACCTTGGATGATGTCAGAGGCTGCCCTTGTGAG
Oy      4201 GATGAGCTTCAAGCTCCCATCATCATCTTTCAGTCAATTTTATTTCTGTTCAAGCTT
Db      4226 GATGAGCTTCAAGCTCCCATCATCATCTTTCAGTCAATTTTATTTCTGTTCAAGCTT
Oy      4261 CTGTACTTTATGAACTTGGCTGTGA
Db      4286 CTGTACTTTATGAACTTGGCTGTGA

```

## RESULT 2

ABQ79964 ID ABQ79964 standard; cDNA; 4761 BP.

AC ABQ79964;

DT 23-DEC-2002 (first entry)

DE Human CD109 K1 protein encoding cDNA.

KM CD109; alpha2 macroglobulin; thioester; cerebroprotective; cardiant;  
immunorepressive; haemostatic; anticoagulant; thrombolytic; human;  
cardiovascular; vasotropic; gene therapy; CD109 K1; gene; ss.

OS Homo sapiens.

Key Location/Qualifiers

FT CDS 113..4450

FT /tag= a

FT /product= "CD109 K1"

XX MO200270696-A2.

XX 12-SEP-2002.

XX 07-MAR-2002; 2002WO-CA00292.

XX 07-MAR-2001; 2001US-273814P.

XX (SCHU/) SCHUH A.

XX (SUTR/) SUTHERLAND R D.

XX Schuh A, Sutherland RD;

XX WPI: 2002-713450/77.

XX P-PSDB: ABB82165.

XX New CD109 nucleic acids and polypeptides, useful in gene therapy,

XX particularly for treating strokes, myocardial infarctions, thrombosis,

XX thrombocytopenia, autoimmune diseases, or organ or bone marrow

XX transplantation -

XX Claim 1; Fig 1a; 156pp; English.

XX The invention relates to isolated nucleic acid molecules encoding CD109

CC polypeptides. These nucleic acid molecules include the human cDNA  
CC sequences comprising CD109 K1, CD109 K1-H7, CD109 K1S or their variants.  
CC CD109 is a novel member of the alpha2 macroglobulin (alpha2m)/C3, C4, C5  
CC family of thioester-containing proteins. The CD109 polypeptides can be  
CC expressed by standard recombinant methodology. The CD109 nucleic acid,  
CC CD109 polypeptide or its mimetic, or the anti-CD109 antibody is useful  
CC for treating or detecting a disease or disorder, e.g. conditions  
CC associated with endothelial activation, platelet activation, activation  
CC of the coagulation or fibrinolytic systems, activation of T lymphocytes  
CC and of the complement system, quantitative or qualitative abnormalities  
CC of platelet function, increased or impaired platelet aggregation and  
CC activation, increased or impaired activation of the coagulation and/or  
CC fibrinolytic systems, or impaired or increased immune activation. These  
CC are also useful for treating cardiovascular disorders, stroke, myocardial  
CC infarction, thrombosis, embolism, peripheral vascular disease,  
CC thrombocytopenia, thrombocythemia, autoimmune diseases, organ  
CC transplantation, or bone marrow transplantation. The CD109 nucleic acids  
CC or their antisense nucleotide sequence are useful in gene therapy for  
CC treating these conditions. The present sequence represents the human  
CC CD109 K1 cDNA sequence.

XX Sequence 4761 BP; 1420 A; 912 C; 1016 G; 1413 T; 0 other;

Query Match 98.5%; Score 4224.4; DB 24; Length 4761;  
Best Local Similarity 98.8%; Pred. No. 0;  
Matches 4286; Conservative 0; Mismatches 1; Indels 51; Gaps 1;

```

Oy      1 ATGACAGGGCCCAACGCTCCGACCGCGCCACCTCTCGGTGTGCAACCGCGGCGTG
Db      113 ATGACAGGGCCCAACGCTCCGACCGCGCCACCTCTCTCGGTGTGCAACCGCGGCGTG
Oy      61 GCCGTGAGCTCCCGGGGCTCGGTTCTGTGACAGCCCGGAGATCATCAGCCCGGAGGA
Db      173 GCCGTGAGCTCCCGGGGCTCGGTTCTGTGACAGCCCGGAGATCATCAGCCCGGAGGA
Oy      121 AATGTACTATTGGGGTGAAGCTTCTGAAACACTGCGCTTACAGGTGACTGTGAAGGGG
Db      233 AATGTACTATTGGGGTGAAGCTTCTGAAACACTGCGCTTACAGGTGACTGTGAAGGGG
Oy      181 GAGCTGCTCAAGACACATCAAAACCTCAGTCTCTGCTGGAACAGAGAGCTTT
Db      293 GAGCTGCTCAAGACACATCAAAACCTCAGTCTCTGCTGGAACAGAGAGCTTT
Oy      241 GAAAAAGGCTCTTTAAGACACTTACTCTTCACTACTCTGTAACAGTGCAGATGAG
Db      353 GAAAAAGGCTCTTTAAGACACTTACTCTTCACTACTCTGTAACAGTGCAGATGAG
Oy      301 ATTATGAGCTACGTGTAAACCGGACGTACCCAGATGAGATTTTATTTCTTAATGATACC
Db      413 ATTATGAGCTACGTGTAAACCGGACGTACCCAGATGAGATTTTATTTCTTAATGATACC
Oy      361 CGCTTATCATTTAGACCAAGAGAAATATGCTTCAATCAACACAAAGGCTTATAC
Db      473 CGCTTATCATTTAGACCAAGAGAAATATGCTTCAATCAACACAAAGGCTTATAC
Oy      421 AAGCCAAAGAGAAATGAAAGTTTCGATTTTACACTCTTCTCAGATTTTAAAGCTTAC
Db      533 AAGCCAAAGAGAAATGAAAGTTTCGATTTTACACTCTTCTCAGATTTTAAAGCTTAC
Oy      481 AAAAAGCTTTTAAACATTTCTCATTAAGAACCCCAATCAAAATTTGATCCACAGTGTG
Db      593 AAAAAGCTTTTAAACATTTCTCATTAAGAACCCCAATCAAAATTTGATCCACAGTGTG
Oy      541 TCACAAAGAGATCTTGAAGATTCATTCACAAATTTTACAGTATCTCCATCCCAATA
Db      653 TCACAAAGAGATCTTGAAGATTCATTCACAAATTTTACAGTATCTCCATCCCAATA
Oy      601 CTGGTGAAGTGTCTATTCAGATTTCAAGTGAAGTGAACAGACATATATCAATATTAC
Db      713 CTGGTGAAGTGTCTATTCAGATTTCAAGTGAAGTGAACAGACATATATCAATATTAC
Oy      661 GTTTCAGAAATATGATTAACCAAAATTTGAAGTGAAGTGAACAGACATATATTTCTT

```

Db 773 GTTTCAGAAATATGTATTAACAAAATTTGAAGTACTTTGCAGACACCATTAATTTGTTCT 832  
Qy 721 ATGAATTCCTAAGCAATTAATAATGTAACATCAAGGCAAAAGTATATATGGAAGCCAGTG 780  
Db 833 ATGAATTCCTAAGCAATTAATAATGTAACATCAAGGCAAAAGTATATATGGAAGCCAGTG 892  
Qy 781 AAAGAGACGTACGCTTACATTTTACCTTTATCTTTTGGGGAAGAAGAAAATATTT 840  
Db 893 AAAGAGACGTACGCTTACATTTTACCTTTATCTTTTGGGGAAGAAGAAAATATTT 952  
Qy 841 ACAAACCAATTAAGATTAATGATCTGCAAACTCTCTTTAATGTAAGAAGATGAA 900  
Db 953 ACAAACCAATTAAGATTAATGATCTGCAAACTCTCTTTAATGTAAGAAGATGAA 1012  
Qy 901 AATGTAATGATTTCTTCAATGAGACTTTCTGAATACCTGATCTATCTTCCCTGGAACA 960  
Db 1013 AATGTAATGATTTCTTCAATGAGACTTTCTGAATACCTGATCTATCTTCCCTGGAACA 1072  
Qy 961 GTAGAAATTTTAAACAAGTGAACAGATCAAGTTAAGATTTTCAAGAAATGTAAGACT 1020  
Db 1073 GTAGAAATTTTAAACAAGTGAACAGATCAAGTTAAGATTTTCAAGAAATGTAAGACT 1132  
Qy 1021 AATGTCCTTCAACCAACAT 1080  
Db 1133 AATGTCCTTCAACCAACAT 1192  
Qy 1081 AAGCATCTCTCAACTTCAACAGCCACTGTGAAGGTATCTGTGATGGAACCAACTG 1140  
Db 1193 AAGCATCTCTCAACTTCAACAGCCACTGTGAAGGTATCTGTGATGGAACCAACTG 1252  
Qy 1141 ACTCTTGAAGAAAGAAAT 1200  
Db 1253 ACTCTTGAAGAAAGAAAT 1312  
Qy 1201 TACTGAGACGGATCTTAACGTGGAATTCAGAAATATGGAAGCTGTTCAGAAAATATAT 1260  
Db 1313 TACTGAGACGGATCTTAACGTGGAATTCAGAAATATGGAAGCTGTTCAGAAAATATAT 1372  
Qy 1261 ACTGTCCCCCAAGTGAACCTTTTAAGATGTAATTCCTGATGGAAGATTCAGAGTGAG 1320  
Db 1373 ACTGTCCCCCAAGTGAACCTTTTAAGATGTAATTCCTGATGGAAGATTCAGAGTGAG 1432  
Qy 1321 CTACAGTTGAAGGCTATTTCTTGTGATGAATAAGATGATGAGCACTTATATGCTGTTT 1380  
Db 1433 CTACAGTTGAAGGCTATTTCTTGTGATGAATAAGATGATGAGCACTTATATGCTGTTT 1492  
Qy 1381 AAGTCTCTAGTAAGACATATCTCAACTAATAACAAGATGAATAATATGAAGGTGGA 1440  
Db 1493 AAGTCTCTAGTAAGACATATCTCAACTAATAACAAGATGAATAATATGAAGGTGGA 1552  
Qy 1441 TCGCCTTTTGAAGTGTGTGATGAGCAACAAGATGTAAGATGTAAGATGTAAGATGTA 1500  
Db 1553 TCGCCTTTTGAAGTGTGTGATGAGCAACAAGATGTAAGATGTAAGATGTAAGATGTA 1612  
Qy 1501 GTATCAGAGGAGCAGTTGGTGTGATGAGCAACAATTTCAACAATGTTCTCTTTAACA 1560  
Db 1613 GTATCAGAGGAGCAGTTGGTGTGATGAGCAACAATTTCAACAATGTTCTCTTTAACA 1672  
Qy 1561 CCAGAAATTTCTTGAAGCTCAAAAGCTGTGTATTTGTATATATGTAAGATGATGGG 1620  
Db 1673 CCAGAAATTTCTTGAAGCTCAAAAGCTGTGTATTTGTATATATGTAAGATGATGGG 1732  
Qy 1621 GAAATTAATGATGATGTTCTAATAATTCCTGTTCAGCTGTTTAAAAATGAATAAG 1680  
Db 1733 GAAATTAATGATGATGTTCTAATAATTCCTGTTCAGCTGTTTAAAAATGAATAAG 1792  
Qy 1681 CTATATTTGAGATGAAGTGAAGCTGAACCATCTGAAGAGTCTCTTGAAGATCTGTGG 1740  
Db 1793 CTATATTTGAGATGAAGTGAAGCTGAACCATCTGAAGAGTCTCTTGAAGATCTGTGG 1852  
Qy 1741 ACAAGCCTGACTCATATGTTGGATTTGAGCTGTGTAACAAGATGTAAGTGTATGAT 1800  
Db 1853 ACAAGCCTGACTCATATGTTGGATTTGAGCTGTGTAACAAGATGTAAGTGTATGAT 1912

Qy 1801 GCCTTAATGATATTTACATGGAATAATGTGTCCATGATGGAACCTTTATTAACAAGGA 1860  
Db 1913 GCCTTAATGATATTTACATGGAATAATGTGTCCATGATGTTGAACCTTTATTAACAAGGA 1972  
Qy 1861 TATTAATTTAGGATGTTTATGATATTTCTTTGCACTTTTCAAGAAATGTGACCTGGGTA 1920  
Db 1973 TATTAATTTAGGATGTTTATGATATTTCTTTGCACTTTTCAAGAAATGTGACCTGGGTA 2032  
Qy 1921 TTGACAGATGCAAACTCAGAGAGGATATATATGATGTTTATGACAAATGCAAGATAT 1980  
Db 2033 TTGACAGATGCAAACTCAGAGAGGATATATATGATGTTTATGACAAATGCAAGATAT 2092  
Qy 1981 GCTGAGATGTTATGAGGAATAATGAGACATATTTGATATATTCATGACTTTTCTTTG 2040  
Db 2093 GCTGAGATGTTATGAGGAATAATGAGACATATTTGATATATTCATGACTTTTCTTTG 2152  
Qy 2041 GGTAGCACTCCACATGTCCGAAGAGATTTCCAGAGACTTGGATTTGGCTAGACCAAC 2100  
Db 2153 GGTAGCACTCCACATGTCCGAAGAGATTTCCAGAGACTTGGATTTGGCTAGACCAAC 2212  
Qy 2101 ATGGTTACAGAAATTTACAGAAATTTGAAGTAACTGTAACCTGATCTATCACTTTGG 2160  
Db 2213 ATGGTTACAGAAATTTACAGAAATTTGAAGTAACTGTAACCTGATCTATCACTTTGG 2272  
Qy 2161 GTGGCTACTGGTTTGTATCTCTGAGGACCTGGGCTTGGACTTAACAACACTACTCAGTG 2220  
Db 2273 GTGGCTACTGGTTTGTATCTCTGAGGACCTGGGCTTGGACTTGAACATCACTCAAGTG 2332  
Qy 2221 GAGCTCAAGCCTTCCAAACATTTTTCATTTTTTGAATCTTCCCTACTCTGTTATCA 2280  
Db 2333 GAGCTCAAGCCTTCCAAACATTTTTCATTTTTTGAATCTTCCCTACTCTGTTATCA 2392  
Qy 2281 GGTGAAGATTTGCTTTGGAATAATATATATATATATATATATATATATATATATAT 2340  
Db 2393 GGTGAAGATTTGCTTTGGAATAATATATATATATATATATATATATATATATATAT 2452  
Qy 2341 AAGGTATCATTTGAGAAAGTGAACAAATTTGATATTTATATGACTTCAAGTGAATAAT 2400  
Db 2453 AAGGTATCATTTGAGAAAGTGAACAAATTTGATATTTATATGACTTCAAGTGAATAAT 2512  
Qy 2401 GCCACAGGCCACAGCAGACACCTTCTGTGTTCCAGTGAAGATGGGGCAACTGTTCTTTT 2460  
Db 2513 GCCACAGGCCACAGCAGACACCTTCTGTGTTCCAGTGAAGATGGGGCAACTGTTCTTTT 2572  
Qy 2461 CCCATCAGGCCAACAACATCTGGAGAAATTCCTATCACAGTCAAGCTTTTCAACCACT 2520  
Db 2573 CCCATCAGGCCAACAACATCTGGAGAAATTCCTATCACAGTCAAGCTTTTCAACCACT 2632  
Qy 2521 GCTTCTGATGCTGTCCACAGATGATTTTATGAATAAGCTGAAGAAATGAATAATCATAT 2580  
Db 2633 GCTTCTGATGCTGTCCACAGATGATTTTATGAATAAGCTGAAGAAATGAATAATCATAT 2692  
Qy 2581 TCACATCATCTTATTTAGACTGACATAGAGGTACAGAGTACCCTGAAAACTTTG 2640  
Db 2693 TCACATCATCTTATTTAGACTGACATAGAGGTACAGAGTACCCTGAAAACTTTG 2752  
Qy 2641 AGTTTCTATTTCTCTCTAATACAGTGACTGGCAGTGAAGAGTTCAATCACTGCAAT 2700  
Db 2753 AGTTTCTATTTCTCTCTAATACAGTGACTGGCAGTGAAGAGTTCAATCACTGCAAT 2812  
Qy 2701 GGAATGTTCTTGGTCTTCCATCAATAGGCTTACCTCATTTGATTCGATTCGATTCG 2760  
Db 2813 GGAATGTTCTTGGTCTTCCATCAATAGGCTTACCTCATTTGATTCGATTCGATTCG 2872  
Qy 2761 TGTGTGAACAGAACATGATTAATTTGTGCTCAATATTTTCACTTTTGGATTAATCTGACT 2820  
Db 2873 TGTGTGAACAGAACATGATTAATTTGTGCTCAATATTTTCACTTTTGGATTAATCTGACT 2932  
Qy 2821 AAAAAAGAAACAACAGTGAAGATTTTGAAGAAAGAGCTTTTCACTTTATGAAGCAAGT 2880  
Db 2933 AAAAAAGAAACAACAGTGAAGATTTTGAAGAAAGAGCTTTTCACTTTATGAAGCAAGT 2992

2881 TACCAGAGAGAACTTCTCTATCAGAGGAGATGAGCTCTTCACTGCTTTGGGAATTAAT 2940  
 2893 TACCAGAGAGAACTTCTCTATCAGAGGAGATGAGCTCTTCACTGCTTTGGGAATTAAT 3052  
 2941 GACCTCTGGAGACCTTGGTGTGACGCTTTGTTTAAAGATGTTTCCCTGAAGCCGAT 3000  
 3053 GACCTCTGGAGACCTTGGTGTGACGCTTTGTTTAAAGATGTTTCCCTGAAGCCGAT 3112  
 3001 CCTTACATAGATATTTGATTCAGAAATGTGTTTACACAGAAATATCATCTGGCTTAAAGACAT 3060  
 3113 CCTTACATAGATATTTGATTCAGAAATGTGTTTACAGAAATATCATCTGGCTTAAAGACAT 3172  
 3061 CAGAAATCCAAAGGCTGAAATTTGGGATCCAGAAAGATGATTCATTAAGTACCTTCAAGT 3120  
 3173 CAGAAATCCAAAGGCTGAAATTTGGGATCCAGAAAGATGATTCATTAAGTACCTTCAAGT 3232  
 3121 GGCATTAAGTCCAGTACAGCTTACAGCCCTATATGTAATTTCTCTCCGGGATATAGA 3180  
 3233 GGCATTAAGTCCAGTACAGCTTACAGCCCTATATGTAATTTCTCTCCGGGATATAGA 3292  
 3181 AAGTATCAGCTTACATGATGTGCAAGAGCTATCCATTTTGGAGTGAATTCAGT 3240  
 3293 AAGTATCAGCTTACATGATGTGCAAGAGCTATCCATTTTGGAGTGAATTCAGT 3352  
 3241 AGAGAAATTTGACAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3300  
 3353 AGAGAAATTTGACAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3412  
 3301 AGTCTTAAAGGAGAGAAAGCTTTTGAATTTGCTGATCTTGAAGAGAGAGAGAGAGAGAGAG 3360  
 3413 AGTCTTAAAGGAGAGAAAGCTTTGAAATTTGCTGATCTTGAAGAGAGAGAGAGAGAGAGAG 3472  
 3361 ATGCAATTTCTGGGTGTCTCATGAGATCCAACTTTCTGATCTCTGAGCCACGCTCCCTG 3420  
 3473 ATGCAATTTCTGGGTGTCTCATGAGATCCAACTTTCTGATCTCTGAGCCACGCTCCCTG 3532  
 3421 GATAATGAAGTTGACGCTTATGACCTGCTTCAACCTTCTTACAAATTTGACATTTCTGAG 3480  
 3533 GATAATGAAGTTGACGCTTATGACCTGCTTCAACCTTCTTACAAATTTGACATTTCTGAG 3592  
 3481 GGAATTCCTCAATTTAGAGTGTGCTAAGCAGCAGAAAGAAATAGCTGGGTGTTTGCATCT 3540  
 3593 GGAATTCCTCAATTTAGAGTGTGCTAAGCAGCAGAAAGAAATAGCTGGGTGTTTGCATCT 3652  
 3541 ACTCAGATACCACTGTGCTTTAAAGCTCTGCTGAATTTGAGCCCTTAATGATATACA 3600  
 3653 ACTCAGATACCACTGTGCTTTAAAGCTCTGCTGAATTTGAGCCCTTAATGATATACA 3712  
 3601 GAAAGACAATAATTCAGATGACCGTGAACGGGCTTACCTGACCAAGTCTT----- 3651  
 3713 GAAAGACAATAATTCAGATGACCGTGAACGGGCTTACCTGACCAAGTCTT----- 3772  
 3652 -----CTTGCTGTGTGTCACGCCA 3669  
 3773 CTGATTGACACACAACCGCTTACTCTTTCAGACAGCAGAGCTGTGTGTGTCACGCCA 3832  
 3670 ATGGCAATTAATTTCCGCAAAATGTTTGGATTTGCTATTTGTCAGCTCAATGTTGTA 3729  
 3833 ATGGCAATTAATTTCCGCAAAATGTTTGGATTTGCTATTTGTCAGCTCAATGTTGTA 3892  
 3730 TATAATGTAAGGCTTCTGGGTCTTCTAGAGAGCAGAGATCTATCCAAATCAAGAACCC 3789  
 3893 TATAATGTAAGGCTTCTGGGTCTTCTAGAGAGCAGAGATCTATCCAAATCAAGAACCC 3952  
 3790 TTTGATTTAATTTCTGCTGTAAAGAAATTAAGATGATCTCAATCATGTGATTTGAT 3849  
 3953 TTTGATTTAATTTCTGCTGTAAAGAAATTAAGATGATCTCAATCATGTGATTTGAT 4012  
 3850 GTGTGTACAAAGCTTTTGGGCGCGGTGAGGTGAGCATGGCTCTTAATGAGAAATTAACCTA 3909  
 4013 GTGTGTACAAAGCTTTTGGGCGCGGTGAGGTGAGCATGGCTCTTAATGAGAAATTAACCTA 4072  
 3910 TTAAGTGGCTTTATGTGCTTTCAGAAAGCAATTTCTCTGACGAGACAGTGAAGAAAGTG 3969

4073 TTAAGTGGCTTTATGTGCTTTCAGAAAGCAATTTCTCTGACGACAGAGAAAGTG 4132  
 3970 GAAATGATCATGAGAAAGCAACCTTATTTGATTTGATTTGATTTGATTTGATTTGATTTGAT 4029  
 4133 GAAATGATCATGAGAAAGCAACCTTATTTGATTTGATTTGATTTGATTTGATTTGATTTGAT 4192  
 4030 GTTAATATTCCTGCTGTGAGAACTTTAAAGTTTCAATTAATCCAGATGCTTCAAGTCTC 4089  
 4193 GTTAATATTCCTGCTGTGAGAACTTTAAAGTTTCAATTAATCCAGATGCTTCAAGTCTC 4252  
 4090 ATAGTGAATTAATTAATGAGCCAGAGACAGGCGGTGAGAAATTAATTAATTAATTAATTAAT 4149  
 4253 ATAGTGAATTAATTAATGAGCCAGAGACAGGCGGTGAGAAATTAATTAATTAATTAATTAAT 4312  
 4150 CTGTCTCTCTGTGACCTTTGACAGTGAATGTCAGAGGCTGCGCTTGTGAGATGAGCT 4209  
 4313 CTGTCTCTCTGTGACCTTTGACAGTGAATGTCAGAGGCTGCGCTTGTGAGATGAGCT 4372  
 4210 TCAGGCTCCCATGATCATCTTCAATTTTATTTTCTGTTTCAAGCTTCTGATCTTT 4269  
 4373 TCAGGCTCCCATGATCATCTTCAATTTTATTTTCTGTTTCAAGCTTCTGATCTTT 4432  
 4270 ATGGAATTTGGCTGTGA 4287  
 4433 ATGGAATTTGGCTGTGA 4450

RESULT 3  
 AAD49440  
 ID AAD49440 standard; DNA; 5882 BP.  
 XX  
 AC AAD49440;  
 XX  
 DT 24-MAR-2003 (first entry)  
 XX  
 DE Human blood cell surface antigen, CD109 encoding DNA #2.  
 XX  
 KW Human; GPI-anchored TGF-beta1 binding protein; r150 protein; cancer;  
 KW glycosylphosphatidylinositol; transforming growth factor-beta1;  
 KW therapy; blood cell surface antigen; CD109; ds.  
 XX  
 OS Homo sapiens.  
 XX  
 PN W0200285942-A2.  
 XX  
 PD 31-OCT-2002.  
 XX  
 PF 24-APR-2002; 2002WC-CA00560.  
 XX  
 PR 24-APR-2001; 2001US-285713P.  
 PR 14-FEB-2002; 2002US-356163P.  
 PA (UVMC-) UNIV MCGILL.  
 PA Philip A. Tam B;  
 PI WPI; 2003-093100/08.  
 DR  
 DR Novel transforming growth factor (TGF)-beta 1 binding reagent which  
 PT comprises r150 protein which acts as accessory receptor of TGF-beta,  
 PT useful for negatively modulating TGF-beta activity, and thus for  
 PT treating cancer -  
 PT  
 XX  
 PS Disclosure, Fig 17, 127pp; English.  
 PS  
 CC The invention relates to novel transforming growth factor (TGF)-beta1  
 CC binding reagent which comprises a glycosylphosphatidylinositol (GPI)-  
 CC anchored TGF-beta1 binding protein referred to as r150 which acts as  
 CC an accessory receptor of TGF-beta. The invention is used for negatively  
 CC modulating TGF-beta activity, and thus for treating conditions  
 CC characterised by overproduction of TGF-beta, such as cancer. Antisense  
 CC molecules of the invention are used for increasing TGF-beta availability



CC and increase graft success. The present sequence is human blood cell  
CC surface antigen, CD109 encoding DNA. CD109 is a variant of r150 sequence.  
CC Note: This sequence is entered to encode human CD109 protein shown in  
CC figure 17 of the specification, but this does not appear to be the case.  
XX  
SQ Sequence 5882 BP; 1696 A; 1173 C; 1241 G; 1772 T; 0 other;

Query Match 98.5%; Score 4224.4; DB 25; Length 5882;  
Best Local Similarity 98.8%; Pred. No. 0;  
Matches 4286; Conservative 0; Mismatches 1; Indels 51; Gaps 1;

1 ATGCGAGGCGCCACCGCTCTGACCGCGCGCCACCTCTCTGCTGTCGACCGCGCGCTG 60  
113 ATGCGAGGCGCCACCGCTCTGACCGCGCGCCACCTCTCTGCTGTCGACCGCGCGCTG 172  
61 GCCGTGCTCCGCGCGCTCGGTTTCTGTGACAGCCCGAGGATCATCGGCCCGAGGA 120  
173 GCCGTGCTCCGCGCGCTCGGTTTCTGTGACAGCCCGAGGATCATCGGCCCGAGGA 232  
121 AATGTGACTATTGGGGTGAAGCTTCTGAAACATGCGCTTCAAGGTGACTGTGAAGCG 180  
233 AATGTGACTATTGGGGTGAAGCTTCTGAAACATGCGCTTCAAGGTGACTGTGAAGCG 292  
181 GAGCTGCTAAGACAGCATCAAACTCATCTGCTCTGCTGAGAGCAAGAGGCTTT 240  
293 GAGCTGCTAAGACAGCATCAAACTCATCTGCTCTGCTGAGAGCAAGAGGCTTT 352  
241 GAAAAAGGCTCTTTTAAAGACCTTACTCTTCACTCACTCTGTAAGAGTGAAGTGA 300  
353 GAAAAAGGCTCTTTTAAAGACCTTACTCTTCACTCACTCTGTAAGAGTGAAGTGA 412  
301 AATTATGAGCTACGTGTAAACCGGACGTACCCAGAGTGAATTTTATCTCTAATAGTAC 360  
413 AATTATGAGCTACGTGTAAACCGGACGTACCCAGAGTGAATTTTATCTCTAATAGTAC 472  
361 CGCTTATCATTTGAGACCAAGAGATATCTGTCTTCAATCAACAGACAAGGCTTATAC 420  
473 CGCTTATCATTTGAGACCAAGAGATATCTGTCTTCAATCAACAGACAAGGCTTATAC 532  
421 AAGCCAAAGCAAGAGTGAAGTGTGCTGTTTACCTCTGCTGAGATTTTAAAGCTTAC 480  
533 AAGCCAAAGCAAGAGTGAAGTGTGCTGTTTACCTCTGCTGAGATTTTAAAGCTTAC 592  
481 AAAACCTCTTTAAACATTTCTCAATTAAGACCCCAATCAAAATTTGATCCAGAGTGTG 540  
593 AAAACCTCTTTAAACATTTCTCAATTAAGACCCCAATCAAAATTTGATCCAGAGTGTG 652  
541 TCACAAACAAAGTATCTTGAAGTCAATTTCCAAAATTTTCAAGTATCTTCCATCCATA 600  
653 TCACAAACAAAGTATCTTGAAGTCAATTTCCAAAATTTTCAAGTATCTTCCATCCATA 712  
601 CTGTGAGTGTGCTTATTAAGTCAAGTGAAGTGAACCAACATTTATTAATCATTTTACG 660  
713 CTGTGAGTGTGCTTATTAAGTCAAGTGAAGTGAACCAACATTTATTAATCATTTTACG 772  
661 GTTTCAGATATGATATTAACAAATTTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 720  
773 GTTTCAGATATGATATTAACAAATTTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 832  
721 ATGAATTTCAAGCATTTAAATGATACATCACGCAAAAGTATATCATATGGAAGCCAGTG 780  
833 ATGAATTTCAAGCATTTAAATGATACATCACGCAAAAGTATATCATATGGAAGCCAGTG 892  
781 AAGAGAGAGTAAACGCTTACATTTTATCTTTTGGGAAAGAAAGAAATAT 840  
893 AAGAGAGAGTAAACGCTTACATTTTATCTTTTGGGAAAGAAAGAAATAT 952  
841 AAAAAAATTTAAGATATGATCTGCAAACTTCTCTTTTAAATGAGAGATGAA 900  
953 AAAAAAATTTAAGATATGATCTGCAAACTTCTCTTTTAAATGAGAGATGAA 1012  
901 AATGTAATGATTTCTTCAATGAGCTTTGTGAATACCTGATCTATCTTCCCTGAGCA 960

1013 AATGTAATGATTTCTTCAATGAGCTTTGTGAATACCTGATCTATCTTCCCTGAGCA 1072  
961 GTAGAAATTTTAAACAGAGTGAACAGATCAGTTACAGGATATTTCAAGAAATGATAGCACT 1020  
1073 GTAGAAATTTTAAACAGAGTGAACAGATCAGTTACAGGATATTTCAAGAAATGATAGCACT 1132  
1021 AATGTTCTTCAAGCAATGATATTAATCATGATGTTTTTGAATTAATCACTGCTTTG 1080  
1133 AATGTTCTTCAAGCAATGATATTAATCATGATGTTTTTGAATTAATCACTGCTTTG 1192  
1081 AAGCATCTCTCAATTTCAAGCCCATGTAAGGTAATCTGCTGATAGGCAACCACTG 1140  
1193 AAGCATCTCTCAATTTCAAGCCCATGTAAGGTAATCTGCTGATAGGCAACCACTG 1252  
1141 ACTCTTGAAGAAAGAAATTAATGATCTTAAACAGTGAACAGAGAACTTAATCTGAG 1200  
1253 ACTCTTGAAGAAAGAAATTAATGATCTTAAACAGTGAACAGAGAACTTAATCTGAG 1312  
1201 TACTGAGCGGATCTTAACAGTGAATCAAGAAATGAAAGCTGTTCAAGAAATTAAT 1260  
1313 TACTGAGCGGATCTTAACAGTGAATCAAGAAATGAAAGCTGTTCAAGAAATTAAT 1372  
1261 ACTGTCCTTCAAGTGAATCTTTAAGTGAATTTCCATCTGAGAGATTTCCAGTGA 1320  
1373 ACTGTCCTTCAAGTGAATCTTTAAGTGAATTTCCATCTGAGAGATTTCCAGTGA 1432  
1321 CTACAGTTGAAGGCTTATTTCTGCTGATGAAGTGAAGTGAAGTGAAGTGAAGTGA 1380  
1433 CTACAGTTGAAGGCTTATTTCTGCTGATGAAGTGAAGTGAAGTGAAGTGAAGTGA 1492  
1381 AAGTCTCTTGAAGCAATATCAATCCAACTAAACCAAGATGAAGAAATTAATGAAGTGA 1440  
1493 AAGTCTCTTGAAGCAATATCAATCCAACTAAACCAAGATGAAGAAATTAATGAAGTGA 1552  
1441 TGCGCTTTGAGTGTGCTGATGAGCAACAGATTTGAAGAGTGAAGTGAAGTGAAGTGA 1500  
1553 TGCGCTTTGAGTGTGCTGATGAGCAACAGATTTGAAGAGTGAAGTGAAGTGAAGTGA 1612  
1501 GTATCAGAGGAGCAGTGTGCTGATGAAGAAACAAATTTCAACAAATGCTCTTAAACA 1560  
1613 GTATCAGAGGAGCAGTGTGCTGATGAAGAAACAAATTTCAACAAATGCTCTTAAACA 1672  
1561 CCAGAAATTTCTTGAATCCAAAGCTGTGAATTTGATTAATTAATGAAGTGAAGTGA 1620  
1673 CCAGAAATTTCTTGAATCCAAAGCTGTGAATTTGATTAATTAATGAAGTGAAGTGA 1732  
1621 GAAATTAATGATGATGTTCTTAAATTTCTGTGACCTGTTTTTAAATTAATGAAGTGA 1680  
1733 GAAATTAATGATGATGTTCTTAAATTTCTGTGACCTGTTTTTAAATTAATGAAGTGA 1792  
1681 CTATATTTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 1740  
1793 CTATATTTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 1852  
1741 ACACAGCTGATCTCAATGTTGAGTGTGATGCTGTTGCAAAAGTGAAGTGAAGTGAAGTGA 1800  
1853 ACACAGCTGATCTCAATGTTGAGTGTGATGCTGTTGCAAAAGTGAAGTGAAGTGAAGTGA 1912  
1801 GCTCTTAATGATTAATTAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 1860  
1913 GCTCTTAATGATTAATTAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 1972  
1861 TATTAATTTAAGCAGTGTGATGAATTTCTTTGCAAGCTTTCAAGAAATGAGTGAAGTGA 1920  
1973 TATTAATTTAAGCAGTGTGATGAATTTCTTTGCAAGCTTTCAAGAAATGAGTGAAGTGA 2032  
1921 TTGACAGATGCAAACTCAAGAGATTAATTAATGATGTTTGAAGTGAAGTGAAGTGAAGTGA 1980  
2033 TTGACAGATGCAAACTCAAGAGATTAATTAATGATGTTTGAAGTGAAGTGAAGTGAAGTGA 2092  
1981 GCTGAGGTTTATGAGAGAAATGAAGACATATGATGATGATGATGATGATGATGATGATGATG 2040  
2093 GCTGAGGTTTATGAGAGAAATGAAGACATATGATGATGATGATGATGATGATGATGATGATG 2152

QY 2041 GGTAGCAGTCCACATGTCGGAAGCATTTTCCAGAGACTTGGAATTTGGCTAGACACCAAC 2100  
DB 2153 GGTGACAGTCCACATGTCGGAAGCATTTTCCAGAGACTTGGAATTTGGCTAGACACCAAC 2212  
QY 2101 ATGGGTTACAGAGATTTTACCAAGAAATTTGAAGTAACTGTAATCTGATTCATCACTTCTGG 2160  
DB 2213 ATGGGTTACAGAGATTTTACCAAGAAATTTGAAGTAACTGTAATCTGATTCATCACTTCTGG 2272  
QY 2161 GTGGCTACTGGTTTTGATCTCTGAGAGACCTGGGCTTTGAGACTTAACTAACTCCAGTG 2220  
DB 2273 GTGGCTACTGGTTTTGATCTCTGAGAGACCTGGGCTTTGAGACTTAACTAACTCCAGTG 2332  
QY 2221 GAGCTCCAGGCTTCCAAACCATTTTTCATTTTGTGAATCTTCCCTACTCTGTTATCAGA 2280  
DB 2333 GAGCTCCAGGCTTCCAAACCATTTTTCATTTTGTGAATCTTCCCTACTCTGTTATCAGA 2392  
QY 2281 GGTGAAGAAATTTGCTTTGGAATAAATAATTCATTAATTTTGAAGATGCCATGAGGT 2340  
DB 2393 GGTGAAGAAATTTGCTTTGGAATAAATAATTCATTAATTTTGAAGATGCCATGAGGT 2452  
QY 2341 AAGGTAATCATTTGAGAAAAGTGAACAAATTTGATATCTTAATGACTTCAAGTAATTAAT 2400  
DB 2453 AAGGTAATCATTTGAGAAAAGTGAACAAATTTGATATCTTAATGACTTCAAGTAATTAAT 2512  
QY 2401 GCCACAGGCCACAGCAGACACCTTCTGGTCCAGTGAAGATGGGGCAACTGTTCTTTT 2460  
DB 2513 GCCACAGGCCACAGCAGACACCTTCTGGTCCAGTGAAGATGGGGCAACTGTTCTTTT 2572  
QY 2461 CCCATCAGGCCCAACATCTGGAGAAAATTCCTATCAGAGTCACAGCTCTTTCACCACT 2520  
DB 2573 CCCATCAGGCCCAACATCTGGAGAAAATTCCTATCAGAGTCACAGCTCTTTCACCACT 2632  
QY 2521 GCTTCTGATGCTGTGCAACCAATGATTTTAAAGGCTGAAGAAATGAAAATTCATAT 2580  
DB 2633 GCTTCTGATGCTGTGCAACCAATGATTTTAAAGGCTGAAGAAATGAAAATTCATAT 2692  
QY 2581 TCACAAATCATCTTAATTAAGACTTGACTGACAAATAGGCTACAGAGTAACTGTAATCTTGG 2640  
DB 2693 TCACAAATCATCTTAATTAAGACTTGACTGACAAATAGGCTACAGAGTAACTGTAATCTTGG 2752  
QY 2641 AGTTTCTCATTTCTCTTAATACAGTGAATGGCAGTGAAGAGTTTCAGATCATGCAAT 2700  
DB 2753 AGTTTCTCATTTCTCTTAATACAGTGAATGGCAGTGAAGAGTTTCAGATCATGCAAT 2812  
QY 2701 GGAATGTTCTTGGTCTTCCATCAATGGCTTAAGCTTATGATTTCCGATGCTTAATGGC 2760  
DB 2813 GGAATGTTCTTGGTCTTCCATCAATGGCTTAAGCTTATGATTTCCGATGCTTAATGGC 2872  
QY 2761 TGTGGTGAACAGAACATGATAATTTTGTCCAAATATTTACATTTTGGATTTATCTGACT 2820  
DB 2873 TGTGGTGAACAGAACATGATAATTTTGTCCAAATATTTACATTTTGGATTTATCTGACT 2932  
QY 2821 AAAAAAGAAACAACCTGACAGATAAATTTGAAGAAAAAGCTCTTTCATTTATGAGCAAGT 2880  
DB 2933 AAAAAAGAAACAACCTGACAGATAAATTTGAAGAAAAAGCTCTTTCATTTATGAGCAAGT 2992  
QY 2881 TACCAAGAGAACTTCTCTATCAGAGGGAATGAGCTTTTCAGTCTTTTGGGAATTAAT 2940  
DB 2993 TACCAAGAGAACTTCTCTATCAGAGGGAATGAGCTTTTCAGTCTTTTGGGAATTAAT 3052  
QY 2941 GACCTTCTGGAGACTTGGTGTGACGCTTTTGTAAAGATGTTTCCCTGAAGCCGAT 3000  
DB 3053 GACCTTCTGGAGACTTGGTGTGACGCTTTTGTAAAGATGTTTCCCTGAAGCCGAT 3112  
QY 3001 CCTTACATAGATATTTGATCAGAAATGTTTACAGAAACATACACTTGCTTAAAGACAT 3060  
DB 3113 CCTTACATAGATATTTGATCAGAAATGTTTACAGAAACATACACTTGCTTAAAGACAT 3172  
QY 3061 CAGAAATCCAAAGGTGAATTTTGGATCCAGAAAGAGTGAATTCATGAGCTTCAAGT 3120  
DB 3173 CAGAAATCCAAAGGTGAATTTTGGATCCAGAAAGAGTGAATTCATGAGCTTCAAGT 3232

QY 3121 GGCATTAAGAGTCCAGTAACACTTACAGCCTTAATTTGTAATCTTCTCTGGGATATAGA 3180  
DB 3233 GGCATTAAGAGTCCAGTAACACTTACAGCCTTAATTTGTAATCTTCTCTGGGATATAGA 3292  
QY 3181 AAGTATACGCTTAACATTTATGATGTGCAAGAGTCTATCCATTTTGTGAGTCTGAATTCAGT 3240  
DB 3293 AAGTATACGCTTAACATTTATGATGTGCAAGAGTCTATCCATTTTGTGAGTCTGAATTCAGT 3352  
QY 3241 AGAGAAATTTACAGCAATTAATCTAGCCTTATTAATCTTATGCAATTTGATCATGAGTGGG 3300  
DB 3353 AGAGAAATTTACAGCAATTAATCTAGCCTTATTAATCTTATGCAATTTGATCATGAGTGGG 3412  
QY 3301 AGTCTTAAGCCAGAGAAAGCTTTGAATATGCTGAATTTGAGAGCAGAAAGAGTGGC 3360  
DB 3413 AGTCTTAAGCCAGAGAAAGCTTTGAATATGCTGAATTTGAGAGCAGAAAGAGTGGC 3472  
QY 3361 ATGCAATTTCTGGGTGTATCAGAGTCCAACTTTTGTGACTCTGGAGCCACGCTCCCTG 3420  
DB 3473 ATGCAATTTCTGGGTGTATCAGAGTCCAACTTTTGTGACTCTGGAGCCACGCTCCCTG 3532  
QY 3421 GATATGAAGTGAAGCCTTATGCACTGCTCAACTTTTACAAATTTGAGACTTCTGAG 3480  
DB 3533 GATATGAAGTGAAGCCTTATGCACTGCTCAACTTTTACAAATTTGAGACTTCTGAG 3592  
QY 3481 GGAATCCCAATTAATGAGTGGCTTAAGCAGCAGCAAGAAATAGCTTGGGTGTTTGCATCT 3540  
DB 3593 GGAATCCCAATTAATGAGTGGCTTAAGCAGCAGCAAGAAATAGCTTGGGTGTTTGCATCT 3652  
QY 3541 ACTCAGATACCACTGTGGCTTTAAAGGCTCTGTCTGAATTTGCAAGCCCTTAATGAATACA 3600  
DB 3653 ACTCAGATACCACTGTGGCTTTAAAGGCTCTGTCTGAATTTGCAAGCCCTTAATGAATACA 3712  
QY 3601 GAAAGGCAATATTCAGAGTGAACCGTGAACGGGCTTACTCAGACAGTCTCTCTCTCTCT 3651  
DB 3713 GAAAGGCAATATTCAGAGTGAACCGTGAACGGGCTTACTCAGACAGTCTCTCTCTCTCTCT 3772  
QY 3652 -----CTTGTGTGTACAGCCA 3669  
DB 3773 CTGATTTGACACACAAACCGCTTACTCTCTCAGACAGCAGCTTGTGTGTGTACAGCCA 3832  
QY 3670 ATGGCAGTTAATTTTCCGCAATGCTTTTGTGATTTGCTAGCTCAATGTTGTA 3729  
DB 3833 ATGGCAGTTAATTTTCCGCAATGCTTTTGTGATTTGCTAGCTCAATGTTGTA 3892  
QY 3730 TATTAAGTGAAGCTTCTGGGCTTCTGAGAAAGCAAGATCTATCCAAATCAAGAGCC 3789  
DB 3893 TATTAAGTGAAGCTTCTGGGCTTCTGAGAAAGCAAGATCTATCCAAATCAAGAGCC 3952  
QY 3790 TTTGATTTAGATGTTGCTGTAAAGAAATGAATGATCTCAATCATGTGATTTGAAT 3849  
DB 3953 TTTGATTTAGATGTTGCTGTAAAGAAATGAATGATCTCAATCATGTGATTTGAAT 4012  
QY 3850 GTGTGTACAGCTTTTGGGCGCGGGTGAAGTGGCATGGCTTTTATGAAATTAACCTA 3909  
DB 4013 GTGTGTACAGCTTTTGGGCGCGGGTGAAGTGGCATGGCTTTTATGAAATTAACCTA 4072  
QY 3910 TTTAAGGCTTTAATGAGCTTCAAGAAAGCAATTTCTCTGAGGAGACAGTGAAGAAAGT 3969  
DB 4073 TTTAAGGCTTTAATGAGCTTCAAGAAAGCAATTTCTCTGAGGAGACAGTGAAGAAAGT 4132  
QY 3970 GAATATGATCATGGAAGAACTCAACTATATTTAGATTTCTGTAAATGAATCCAGTTTGT 4029  
DB 4133 GAATATGATCATGGAAGAACTCAACTATATTTAGATTTCTGTAAATGAATCCAGTTTGT 4192  
QY 4030 GTTAATATTTCTGCTGTGAGAAACTTTAAAGTTTCAATATCCCAAGATGCTTCAGTCTC 4089  
DB 4193 GTTAATATTTCTGCTGTGAGAAACTTTAAAGTTTCAATATCCCAAGATGCTTCAGTCTC 4252  
QY 4090 ATAGTGAATTAATGAGCCAAAGAGACAGGCGGTGAGAAATTAACAATGAGTGAAG 4149  
DB 4253 ATAGTGAATTAATGAGCCAAAGAGACAGGCGGTGAGAAATTAACAATGAGTGAAG 4312  
QY 4150 CTGTCTCTGTGACCTTTGCACTGATGTCAGAGGCTGCGTCTTGTGAGATGAGCT 4209

Db 4313 CTGCTCTCTGACCTTTCAGTGAATCCAGGAGCTGCTCTTGTAGAGATGAGCT 4372  
 QY 4210 TCAGGCTCCCATCATCATCTTCAGATCTTATTTTCGTTCAAGCTTCGACTT 4269  
 Db 4373 TCAGGCTCCCATCATCATCTTCAGATCTTATTTTCGTTCAAGCTTCGACTT 4432  
 QY 4270 ATGGAACCTTGGCTGTGA 4287  
 Db 4433 ATGGAACCTTGGCTGTGA 4450  
 RESULT 4  
 ABQ79966 standard; cDNA; 5895 BP.  
 ID ABQ79966  
 XX  
 AC ABQ79966;  
 XX  
 DT 23-DEC-2002 (first entry)  
 XX  
 DE Human CD109 KI-H7 protein encoding cDNA.  
 XX  
 KM CD109; alpha2 macroglobulin; thioester; cerebroprotective; cardiant;  
 KM immunosuppressive; haemostatic; anticoagulant; thrombolytic; human;  
 KM cardiovascular; vasotropic; gene therapy; CD109 KI-H7; gene; ss.  
 OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 FT CDS 113..4450  
 FT /tag= a  
 FT /product= "CD109 KI-H7"  
 PN MO200270696-A2.  
 XX  
 PD 12-SEP-2002.  
 XX  
 PF 07-MAR-2002; 2002WC-CA00292.  
 XX  
 PR 07-MAR-2001; 2001US-273814P.  
 XX  
 PA (SCHU/) SCHUH A.  
 PA (SUTH/) SUTHERLAND R D.  
 PI Schuh A, Sutherland RD;  
 XX  
 DR WPI; 2002-713450/77.  
 DR P-PSDB; ABB82167.  
 XX  
 PT New CD109 nucleic acids and polypeptides, useful in gene therapy,  
 PT particularly for treating strokes, myocardial infarctions, thrombosis,  
 PT thrombocytopenia, autoimmune diseases, or organ or bone marrow  
 PT transplantation.  
 PS Claim 1; Fig 2a; 156pp; English.  
 XX  
 CC The invention relates to isolated nucleic acid molecules encoding CD109  
 CC polypeptides. These nucleic acid molecules include the human CDNA  
 CC sequences comprising CD109 KI, CD109 KI-H7, CD109 K15 or their variants.  
 CC CD109 is a novel member of the alpha2 macroglobulin (alpha2M)/C3, C4, C5  
 CC family of thioester-containing proteins. The CD109 polypeptides can be  
 CC expressed by standard recombinant methodology. The CD109 nucleic acid,  
 CC CD109 polypeptide or its mimetic, or the anti-CD109 antibody is useful  
 CC for treating or detecting a disease or disorder, e.g. conditions  
 CC associated with endothelial activation, platelet activation, activation  
 CC of the coagulation or fibrinolytic systems, activation of T lymphocytes  
 CC and of the complement system, quantitative or qualitative abnormalities  
 CC of platelet function, increased or impaired platelet aggregation and/or  
 CC activation, increased or impaired activation of the coagulation and/or  
 CC fibrinolytic systems, or impaired or increased immune activation. These  
 CC are also useful for treating cardiovascular disorders, stroke, myocardial  
 CC infarction, thrombosis, embolism, peripheral vascular disease,  
 CC thrombocytopenia, thrombocytopenia, autoimmune diseases, organ

CC transplantation, or bone marrow transplantation. The CD109 nucleic acids  
 CC or their antisense nucleotide sequence are useful in gene therapy for  
 CC treating these conditions. The present sequence represents the human  
 CC CD109 KI-H7 cDNA sequence.  
 XX  
 SQ Sequence 5895 BP; 1709 A; 1173 C; 1241 G; 1772 T; 0 other;  
 Query Match 98.5%; Score 4224.4; DB 24; Length 5895;  
 Best Local Similarity 98.8%; Pred. No. 0;  
 Matches 4286; Conservative 0; Mismatches 1; Indels 51; Gaps 1;  
 QY 1 ATGCAAGGCGCCACCGCTCTGACCGCGCCACCTCTCTGCTGTCACCGCGCTG 60  
 Db 113 ATGCAAGGCGCCACCGCTCTGACCGCGCCACCTCTCTGCTGTCACCGCGCTG 172  
 QY 61 GCCGTGGCTCCCGGGCCCTGGTTCTGGTGAACGCCAGGAGATCATAGGCCGGAGA 120  
 Db 173 GCCGTGGCTCCCGGGCCCTGGTTCTGGTGAACGCCAGGAGATCATAGGCCGGAGA 232  
 QY 121 AATGTACTAATYGGGGTGAAGCTTCTGGACAATGCCCTTTCACAGTGAAGTGAAGCG 180  
 Db 233 AATGTACTAATYGGGGTGAAGCTTCTGGACAATGCCCTTTCACAGTGAAGTGAAGCG 292  
 QY 181 GAGCTGCTCAAGACAGATCAAACTCTGCTCTGCTGGAAGAGAGAGCTTTT 240  
 Db 293 GAGCTGCTCAAGACAGATCAAACTCTGCTCTGCTGGAAGAGAGAGCTTTT 352  
 QY 241 GAAAAAGGCTCTTTTAAACACTTAATCTTCTCATCACTCACTGGAAGAGAGATGAG 300  
 Db 353 GAAAAAGGCTCTTTTAAACACTTAATCTTCTCATCACTCACTGGAAGAGAGATGAG 412  
 QY 301 ATTTATGACTACGTTGTAACCGGACGTACCGGAGTGAATTTTATCTTAATAGTACC 360  
 Db 413 ATTTATGACTACGTTGTAACCGGACGTACCGGAGTGAATTTTATCTTAATAGTACC 472  
 QY 361 CGCTTATCATTTAGACACCAAGAGATATCTGCTTCAATCAAGACCAAGGCTTATAC 420  
 Db 473 CGCTTATCATTTAGACACCAAGAGATATCTGCTTCAATCAAGACCAAGGCTTATAC 532  
 QY 421 AAGCCAAAGCAAGAGTAAGTGAAGTGTGATGATGATGATGATGATGATGATGATG 480  
 Db 533 AAGCCAAAGCAAGAGTAAGTGAAGTGTGATGATGATGATGATGATGATGATGATG 592  
 QY 481 AAAACCTCTTAAACATTTCTAATTAAGGACCCCAATCAATTTGATGATCAAGGTG 540  
 Db 593 AAAACCTCTTAAACATTTCTAATTAAGGACCCCAATCAATTTGATGATCAAGGTG 652  
 QY 541 TCACAAACAAGTGAATCTTGAGTCAATTTCCAAAATTTTCAGCTATCTCCATCCATA 600  
 Db 653 TCACAAACAAGTGAATCTTGAGTCAATTTCCAAAATTTTCAGCTATCTCCATCCATA 712  
 QY 601 CTGGTGAATGCTATTCAGAGTCAAGTGAATGACGACATATTAATCAATTCAG 660  
 Db 713 CTGGTGAATGCTATTCAGAGTCAAGTGAATGACGACATATTAATCAATTCAG 772  
 QY 661 GTTTCAGATATGATTAACCAAAATTTGAATGATTTGAGACACCAATTAATGTTCT 720  
 Db 773 GTTTCAGATATGATTAACCAAAATTTGAATGATTTGAGACACCAATTAATGTTCT 832  
 QY 721 ATGAATTTAAGCATTAAATGATGATGATGATGATGATGATGATGATGATGATGATG 780  
 Db 833 ATGAATTTAAGCATTAAATGATGATGATGATGATGATGATGATGATGATGATGATG 892  
 QY 781 AAGGAGACGTAGAGCTTACATTTTACCTTATCTTTGGGGAAGAGAAATATTT 840  
 Db 893 AAGGAGACGTAGAGCTTACATTTTACCTTATCTTTGGGGAAGAGAAATATTT 952  
 QY 841 ACAAATAATTAAGATTAAGATGATGATGATGATGATGATGATGATGATGATGATGATG 900  
 Db 953 ACAAATAATTAAGATTAAGATGATGATGATGATGATGATGATGATGATGATGATGATG 1012  
 QY 901 AATGTAATGATTTCTTCAATGATGATGATGATGATGATGATGATGATGATGATGATG 960

Db	1013	AATGTAATGAGATTCTTCAAATGSACTTTCTGAATACCTGAGATCTATCTTCCCTGAGACA	1072
Qy	961	GTAGAAATTTTAAACCAAGTGAACAGAAATCAGTTACAGGTATTTTCAAGAAATGTAAGCACT	1020
Db	1073	GTAGAAATTTTAAACCAAGTGAACAGAAATCAGTTACAGGTATTTTCAAGAAATGTAAGCACT	1132
Qy	1021	AATGAGTCTTCAACCAACATGATTAACATCAATGAGTTTGTGATTAATCAATCTGTG	1080
Db	1133	AATGAGTCTTCAACCAACATGATTAACATCAATGAGTTTGTGATTAATCAATCTGTG	1192
Qy	1081	AAGCCATCTCAACATTTCAACAGCCACTGTGAAGTAACTGTCGATGAGCAACCACTG	1140
Db	1193	AAGCCATCTCTCAACTTTCAACGCCACTGTGAAGTAACTGTCGATGAGCAACCACTG	1252
Qy	1141	ACTCTTGAAGAAAGAAATATATGTAGTCAATACAGTGAACAGAGAACTATATCTGAG	1200
Db	1253	ACTCTTGAAGAAAGAAATATATGTAGTCAATACAGTGAACAGAGAACTATATCTGAG	1312
Qy	1201	TACTGGAACGGATCTTAACAGTGAATCAAGAAATGGAAGCTGTTCAAGAAATTAATAT	1266
Db	1313	TACTGGAACGGATCTTAACAGTGAATCAAGAAATGGAAGCTGTTCAAGAAATTAATAT	1372
Qy	1261	ACTGCCCCCAAGTGAACCTTTTAAGATTGAATCCCATCCGTGAGATTTCCAGTGA	1320
Db	1373	ACTGCCCCCAAGTGAACCTTTTAAGATTGAATCCCATCCGTGAGATTTCCAGTGA	1432
Qy	1321	CTACAGTTGAAGCCCTATTTCTTGTGTATAAAGTAGCATGCAAGTTCAATAGTCTGTT	1380
Db	1433	CTACAGTTGAAGCCCTATTTCTTGTGTATAAAGTAGCATGCAAGTTCAATAGTCTGTT	1492
Qy	1381	AAGTCTCTAGTAAGACATACATCCAATAAACAGAGATGAATAATTAAGTGGGA	1440
Db	1493	AAGTCTCTAGTAAGACATACATCCAATAAACAGAGATGAATAATTAAGTGGGA	1552
Qy	1441	TGACCTTTGAGTGTGGTATAGTGGCAACAAAGATTGAAGGATTAAGTATATGTTA	1500
Db	1553	TGACCTTTGAGTGTGGTATAGTGGCAACAAAGATTGAAGGATTAAGTATATGTTA	1612
Qy	1501	GTATCCAGGGGACAGTGTGGCTGTAGAAACCAAAATTCACATGTTCTCTTAAACA	1560
Db	1613	GTATCCAGGGGACAGTGTGGCTGTAGAAACCAAAATTCACATGTTCTCTTAAACA	1672
Qy	1561	CCAGAAATTTCTTGACATCCAAAGCCGTGTATGTGTATATTAATGAAGATGATGGG	1620
Db	1673	CCAGAAATTTCTTGACATCCAAAGCCGTGTATGTGTATATTAATGAAGATGATGGG	1732
Qy	1621	GAAATTAATGAAGATGTTCTAATAATTCCTGTCAAGCTGTTTAAATAATGAATGAAG	1680
Db	1733	GAAATTAATGAAGATGTTCTAATAATTCCTGTCAAGCTGTTTAAATAATGAATGAAG	1792
Qy	1681	CTATATTTAGAGTAAAGTGAAGAGCTGAACCATCTGAGAAAGTCTCTAGGATCTGTG	1740
Db	1793	CTATATTTAGAGTAAAGTGAAGAGCTGAACCATCTGAGAAAGTCTCTAGGATCTGTG	1852
Qy	1741	AACAGCCTGACTCCATAGTTGGGATTTGAGCTGTGAACAAAGTGAATCTGATGAT	1800
Db	1853	AACAGCCTGACTCCATAGTTGGGATTTGAGCTGTGAACAAAGTGAATCTGATGAT	1912
Qy	1801	GCCCTTAATGAATATTAACAATGGAATAATGTGCTCAATGATTTGAACCAACAGGA	1860
Db	1913	GCCCTTAATGAATATTAACAATGGAATAATGTGCTCAATGATTTGAACCAACAGGA	1972
Qy	1861	TATTAATTTAGGATTTTATGATTAATCTTTTGAGCTTTCAAGAAATGGAATCTGGGTA	1920
Db	1973	TATTAATTTAGGATTTTATGATTAATCTTTTGAGCTTTCAAGAAATGGAATCTGGGTA	2032
Qy	1921	TTGACAGATGCAAACTTCAGAGAGATTATATGATGTTTATGAACATGCAAGATAT	1980
Db	2033	TTGACAGATGCAAACTTCAGAGAGATTATATGATGTTTATGAACATGCAAGATAT	2092
Qy	1981	GCTGAGAGGTTTATGAGAGAAATGAAGACATATGTGAATATTCATGATCTTTCTTGG	2040
Db	2093	GCTGAGAGGTTTATGAGAGAAATGAAGACATATGTGAATATTCATGATCTTTCTTGG	2152

QY	2041	GGTAGACAGTCCAACTGTGCCGAAGACATTTTCAGAGACTTGAGATTGGCTTAGACACCAAC	2100
Db	2153	GGTAGACAGTCCAACTGTGCCGAAGACATTTTCAGAGACTTGAGATTGGCTTAGACACCAAC	2212
QY	2101	ATGGGTTACAGAGATTTTACCAAGATTTGAAGTAACTGATCCTGATCTTATCACTCTCTTGG	2160
Db	2213	ATGGGTTACAGAGATTTTACCAAGATTTGAAGTAACTGATCCTGATCTTATCACTCTCTTGG	2272
QY	2161	GTGGCTACTGATTTTGTGATCTCTGAGAGCCTGGGCTTGGACTAACACTACTCCAGTG	2220
Db	2273	GTGGCTACTGATTTTGTGATCTCTGAGAGCCTGGGCTTGGACTAACACTACTCCAGTG	2332
QY	2221	GAGCTCCAGGCTTCCAACTTTTCAATTTTTTGAATCTTCCCTACTCTGTATATAGA	2280
Db	2333	GAGCTCCAGGCTTCCAACTTTTCAATTTTTTGAATCTTCCCTACTCTGTATATAGA	2392
QY	2281	GGTAGAAGATTGGTTTGGAAATACATATATTCATATTTGAAGAAGATGCACTGAGGTT	2340
Db	2393	GGTAGAAGATTGGTTTGGAAATACATATATTTCAATATTTGAAGAAGATGCACTGAGGTT	2452
QY	2341	AAGGTAATCATTTGAGAAAAGTGAACAATTTTGATATTTCTAATGACCTTCAGTGAATTAAT	2400
Db	2453	AAGGTAATCATTTGAGAAAAGTGAACAATTTTGATATTTCTAATGACCTTCAGTGAATTAAT	2512
QY	2401	GCCAAGGCCACGACGAGACCTTTCTGTTCCAGTAGAGATGGGGCACTGTTCTTTT	2460
Db	2513	GCCAAGGCCACGACGAGACCTTTCTGTTCCAGTAGAGATGGGGCACTGTTCTTTT	2572
QY	2461	CCCATTCAGGCCAACAACATCTGGGGAATTCCTATGCACTGCACAGTCTTTTCAACCAC	2520
Db	2573	CCCATTCAGGCCAACAACATCTGGGGAATTCCTATGCACTGCACAGTCTTTTCAACCAC	2632
QY	2521	GCTTCTGATGCTGTACCCAGATATATTTAGTAAAGGCTGAAGAAATAGAAAATCATAT	2580
Db	2633	GCTTCTGATGCTGTACCCAGATATATTTAGTAAAGGCTGAAGAAATAGAAAATCATAT	2692
QY	2581	TCACAATTCATCTTATTTAGACTGTGACATATAGGCTACAGAGTACCTGAAAACCTTGG	2640
Db	2693	TCACAATTCATCTTATTTAGACTGTGACATATAGGCTACAGAGTACCTGAAAACCTTGG	2752
QY	2641	AGTTTCTCATTTCTCTCTAATACGTGACTGGCGAGTGAAGAAGTTTCAGATCATGCAATT	2700
Db	2753	AGTTTCTCATTTCTCTCTAATACGTGACTGGCGAGTGAAGAAGTTTCAGATCATGCAATT	2812
QY	2701	GGAGATGTTCTTGGTCCCTCCATCAATGGCTAGCCTCATTTGATTTGGATGCTTATGCT	2760
Db	2813	GGAGATGTTCTTGGTCCCTCCATCAATGGCTAGCCTCATTTGATTTGGATGCTTATGCT	2872
QY	2761	TGTGTGTGAACGAACATGATTAATTTTGTCTCCAAATTTTACATTTTGGATATCTGACT	2820
Db	2873	TGTGTGTGAACGAACATGATTAATTTTGTCTCCAAATTTTACATTTTGGATATCTGACT	2932
QY	2821	AAAAAGAAACAACCTGACAGATTAATTTGAAGAAGAAAGCTCTTTCATTTATGAGCGAAGT	2880
Db	2933	AAAAAGAAACAACCTGACAGATTAATTTGAAGAAGAAAGCTCTTTCATTTATGAGCGAAGT	2992
QY	2881	TACAGAGAGAACTTCTCTATACAGAGGAAGATGGCTCTTCACTGCTGTTTGGGAATTAAT	2940
Db	2993	TACAGAGAGAACTTCTCTATACAGAGGAAGATGGCTCTTCACTGCTGTTTGGGAATTAAT	3052
QY	2941	GACCCCTCTGGAGACACTTGGTGTGACGCTTGTGTTTAAAGATGTTCTTGAAGCCGAT	3000
Db	3053	GACCCCTCTGGAGACACTTGGTGTGACGCTTGTGTTTAAAGATGTTCTTGAAGCCGAT	3112
QY	3001	CCTTACATAGATATTTGATCAGATGTGTTCACAGAAACATACCTTGGCTTTAAAGACAT	3060
Db	3113	CCTTACATAGATATTTGATCAGATGTGTTCACAGAAACATACCTTGGCTTTAAAGACAT	3172
QY	3061	CAGAAATTCACACGGTGAATTTTGGAGTCCAGGAAGATGATTAATGTGAGCTTCAAGGT	3120
Db	3173	CAGAAATTCACACGGTGAATTTTGGAGTCCAGGAAGATGATTAATGTGAGCTTCAAGGT	3232

```

QY 3121 GGCAATTAAGTCAGTAACTTACAGCCTATATTGTAATCTCTCTCTGGGATATAGA 3180
Db 3233 GGCAATTAAGTCAGTAACTTACAGCCTATATTGTAATCTCTCTCTGGGATATAGA 3292
QY 3181 AAGTATCAGCCTTAACATGATGCAAGAGTCAATCATTTTGGAGTCGAATTCAGT 3240
Db 3239 AAGTATCAGCCTTAACATGATGCAAGAGTCAATCATTTTGGAGTCGAATTCAGT 3352
QY 3241 AGAGAAATTTGAGACAAATTATCTAGCCCTTAATACTTATGCAATTCAGTGGAG 3300
Db 3353 AGAGAAATTTGAGACAAATTATCTAGCCCTTAATACTTATGCAATTCAGTGGAG 3412
QY 3301 AGTCTTAAAGCAGAGAGAGCTTTGAATATGCTGACTTGAGAGCAGAAACAAGAGTGC 3360
Db 3413 AGTCTTAAAGCAGAGAGAGCTTTGAATATGCTGACTTGAGAGCAGAAACAAGAGTGC 3472
QY 3361 ATGCAATTCCTGGGTGTCATCAGATCCAACTTTCTGANTCCTGGCAGCAGCTCCCTG 3420
Db 3473 ATGCAATTCCTGGGTGTCATCAGATCCAACTTTCTGANTCCTGGCAGCAGCTCCCTG 3532
QY 3421 GATATTTGAGTTGACAGCTATGCACTGCTCTCAACTTCTTACAATTTGCACTTCTGAG 3480
Db 3533 GATATTTGAGTTGACAGCTATGCACTGCTCTCAACTTCTTACAATTTGCACTTCTGAG 3592
QY 3481 GGAATCCCAATTTATGAGTGGCTTAAGCAGCAAGAAATAGCTTGGTGGTTTTCATCT 3540
Db 3593 GGAATCCCAATTTATGAGTGGCTTAAGCAGCAAGAAATAGCTTGGTGGTTTTCATCT 3652
QY 3541 ACTCAGATATACACATGCTGCTTTAAAGCTCTGTCTGAAATTTGGACCCCTATGAAATACA 3600
Db 3653 ACTCAGATATACACATGCTGCTTTAAAGCTCTGTCTGAAATTTGGACCCCTATGAAATACA 3712
QY 3601 GAAAGGACAAATATCCAGTGAACCGTGAACGGGACCTAGCTCACAAGTCTCT----- 3651
Db 3713 GAAAGGACAAATATCCAGTGAACCGTGAACGGGACCTAGCTCACAAGTCTCTGTAAGTTT 3772
QY 3652 -----CTTCTGTGTGTAGAGCCA 3669
Db 3773 CTGATTTGACACACACACACCGCTTACTCTTCAAGACAGAGCTTGTGTAGAGCCA 3832
QY 3670 ATGGCAGTTAATATTTCCGCAATATGTTTGGATTTGCTATTGTCTCAGTCAATGTTGTA 3729
Db 3833 ATGGCAGTTAATATTTCCGCAATATGTTTGGATTTGCTATTGTCTCAGTCAATGTTGTA 3892
QY 3730 TATTAATGTAAGGCTTTCGGGCTTCTAGAGAACGAGATCTATCCAAATCAAGAGCC 3789
Db 3893 TATTAATGTAAGGCTTTCGGGCTTCTAGAGAACGAGATCTATCCAAATCAAGAGCC 3952
QY 3790 TTTGATTTAGATGTTGCTGTATAAGAAATAAAGATGATCTCAATCATGTGATTTGAAT 3849
Db 3953 TTTGATTTAGATGTTGCTGTATAAGAAATAAAGATGATCTCAATCATGTGATTTGAAT 4012
QY 3850 GTGTGTACAAGCTTTTCGGGCCCCGGGTAGAGATGTCATGCTTATGAGAAATTAACTTA 3909
Db 4013 GTGTGTACAAGCTTTTCGGGCCCCGGGTAGAGATGTCATGCTTATGAGAAATTAACTTA 4072
QY 3910 TTAAGTGGCTTATAGTGTCTTCAAGAGCAATTTCTGAGAGGAGACAGTGAAGAAAGTG 3969
Db 4073 TTAAGTGGCTTATAGTGTCTTCAAGAGCAATTTCTGAGAGGAGACAGTGAAGAAAGTG 4132
QY 3970 GAATATGATCATGAGAACTCAACCTCTATTATGATTTGTAATGAAACCAAGTTTGT 4029
Db 4133 GAATATGATCATGAGAACTCAACCTCTATTATGATTTGTAATGAAACCAAGTTTGT 4192
QY 4030 GTTAATATTTCTGCTGTGAGAAACTTTAAAGTTTCAATACCAAGATGCTTCAAGTCTC 4089
Db 4193 GTTAATATTTCTGCTGTGAGAAACTTTAAAGTTTCAATACCAAGATGCTTCAAGTCTC 4252
QY 4090 AATAGTGAATTAATATGAGCAAGAGAGACAGCGGTGAGAAAGTTAAACAATCTGAATGAAG 4149
Db 4253 AATAGTGAATTAATATGAGCAAGAGAGACAGCGGTGAGAAAGTTAAACAATCTGAATGAAG 4312
QY 4150 CTGTCTCTGTGACCTTTGACAGATATGTCAGGCGTCCGTCTTGTGAGATGAGACT 4209

```

```

Db 4313 CTGTCTCTGTGACCTTTGACAGATGTCAGGCGCTGCCCTTGTGAGATGAGACT 4372
QY 4210 TCAGGCTCCCATCAACACTCTTCACTGATTTTATTTTCTGTTCAGGCTCTGACTTT 4269
Db 4373 TCAGGCTCCCATCAACACTCTTCACTGATTTTATTTTCTGTTCAGGCTCTGACTTT 4432
QY 4270 ATGGAATTTTGGCTGTGA 4287
Db 4433 ATGGAATTTTGGCTGTGA 4450

RESULT 5
AAD49435
ID AAD49435 standard; DNA; 5883 BP.
XX
AC AAD49435;
XX
DT 24-MAR-2003 (first entry)
XX
DE Human r150 DNA #1.
XX
KW Human, GPI-anchored TGF-beta1 binding protein; r150 protein; cancer;
KM glycosylphosphatidylinositol; transforming growth factor-beta1;
XX therapy; gene; ds.
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 113..4450
FT /tag= a
FT /product= "Human r150 protein #2"
FT /transl_except= (pos:2219..2221, aa:taa)
FT /note= "Xaa corresponds to Ser, Tyr"
XX
FN W0200285942-A2.
XX
PD 31-OCT-2002.
XX
PF 24-APR-2002; 2002MO-CA00560.
XX
PR 24-APR-2001; 2001US-285713P.
PR 14-FEB-2002; 2002US-356163P.
XX
XX (UNMC-) UNIV MCGILL.
XX PA Phillip A, Tam B;
XX PI
XX DR WPI; 2003-093100/08.
XX P-PSDB; ABE32013.
XX
PT Novel transforming growth factor (TGF)-beta 1 binding reagent which
PT comprises r150 protein which acts as accessory receptor of TGF-beta,
PT useful for negatively modulating TGF-beta activity, and thus for
PT treating cancer -
XX
XX PS Claim 7; Page 100-103; 127pp; English.
XX
XX CC The invention relates to novel transforming growth factor (TGF)-beta1
XX CC binding reagent which comprises a glycosylphosphatidylinositol (GPI)-
XX CC anchored TGF-beta1 binding protein referred to as r150 which acts as
XX CC an accessory receptor of TGF-beta. The invention is used for negatively
XX CC modulating TGF-beta activity, and thus for treating conditions
XX CC characterised by overproduction of TGF-beta, such as cancer. Antisense
XX CC molecules of the invention are used for increasing TGF-beta availability
XX CC and increase graft success. The present sequence is human r150 protein
XX CC encoding DNA.
XX
SQ Sequence 5883 BP; 1696 A; 1173 C; 1241 G; 1772 T; 1 other;
XX
Query Match 98.5%; Score 4223.4; DB 25; Length 5883;
Best Local Similarity 98.8%; Pred. No. 0;
Matches 4285; Conservative 0; Mismatches 2; Indels 51; Gaps 1;

```



OY	1	ATGCAAGGCGCCACCGCTCTCTGACCGCGCCGCACTTCTCTGTGCTGTGACCGCGCGCTG	60
Db	113	ATGCAAGGCGCCACCGCTCTCTGACCGCGCCGCACTTCTCTGTGCTGTGACCGCGCGCTG	172
OY	61	GCCGGGCGCTCCCGGCGCTCGGTTTTCTGGTGACGCGCCAGGGATCATGACGGCCGGAGGA	120
Db	173	GCCGGGCGCTCCCGGCGCTCGGTTTTCTGGTGACGCGCCAGGGATCATGACGGCCGGAGGA	232
OY	121	AATGTGACTATTTGGGGTGGAGCTTTGTGAAACACTGCGCTTCACAGGTGACTGTGAAGCG	180
Db	233	AATGTGACTATTTGGGGTGGAGCTTTGTGAAACACTGCGCTTCACAGGTGACTGTGAAGCG	292
OY	181	GAGCTGCTCAAGACAGCAATCAAACTCACTGTCTGTGTCCTGGAAGCAAGAGAGCTTTT	240
Db	293	GAGCTGCTCAAGACAGCAATCAAACTCACTGTCTGTGTCCTGGAAGCAAGAGAGCTTTT	352
OY	241	GAAAAAGGCTCTTTTAAGACACTTACTGTCCATACATCACTGTGAACAGTGCGAGTAG	300
Db	353	GAAAAAGGCTCTTTTAAGACACTTACTGTCCATACATCACTGTGAACAGTGCGAGTAG	412
OY	301	ATTATGAGCTACCGTGTAAACGGGACGTACCGAGATGAGATTTTATTTCTAAATGTACC	360
Db	413	ATTATGAGCTACCGTGTAAACGGGACGTACCGAGATGAGATTTTATTTCTAAATGTACC	472
OY	361	CGCTTATCAATTTGAGACCAAGAAATCTGTCTTTCATTCAACAGACAGCGCTTATAC	420
Db	473	CGCTTATCAATTTGAGACCAAGAAATCTGTCTTTCATTCAACAGACAGCGCTTATAC	532
OY	421	AAGCCAAAGCAAGATGTAAGTTTGGATTGTATACCTTCTCAGATTTTAAAGCTTAC	480
Db	533	AAGCCAAAGCAAGATGTAAGTTTGGATTGTATACCTTCTCAGATTTTAAAGCTTAC	592
OY	481	AAAACCTCTTTTAAACATTCCTCATTTAAGAACCCCAATCAAAATTTGATCAACAGGTTG	540
Db	593	AAAACCTCTTTTAAACATTCCTCATTTAAGAACCCCAATCAAAATTTGATCAACAGGTTG	652
OY	541	TCACAACAAGATCTTGGAGTCATTTCCAAACTTTTCAGCTATCTTCCATCCATA	600
Db	653	TCACAACAAGATCTTGGAGTCATTTCCAAACTTTTCAGCTATCTTCCATCCATA	712
OY	601	CTTGTGACTGCTCATTTCAAGTTTCAAGTGAATGACCGAGCATATTTACATCATTTGAG	660
Db	713	CTTGTGACTGCTCATTTCAAGTTTCAAGTGAATGACCGAGCATATTTACATCATTTGAG	772
OY	661	GTTTCAGAAATGTATTTACCAAAATTTGAAGTACTTTCGAGACACCATATATTTGTTCT	720
Db	773	GTTTCAGAAATGTATTTACCAAAATTTGAAGTACTTTCGAGACACCATATATTTGTTCT	832
OY	721	ATGATTTCTAAGCATTTAAATGTTATCATCAACGCGAAAGTATACATATGGGAAGCAGTG	780
Db	833	ATGATTTCTAAGCATTTAAATGTTATCATCAACGCGAAAGTATACATATGGGAAGCAGTG	892
OY	781	AAAGGAGACGTACCGCTTACATTTTACCTTATATCTTTTGGGGAAGAGAAAATATTT	840
Db	893	AAAGGAGACGTACCGCTTACATTTTACCTTATATCTTTTGGGGAAGAGAAAATATTT	952
OY	841	ACAAAACATTTAATATTAATGATCTGCAAACTCTCTTTTAAATGAAGAGATGAATA	900
Db	953	ACAAAACATTTAATATTAATGATCTGCAAACTCTCTTTTAAATGAAGAGATGAATA	1012
OY	901	AATGTAAATGATCTTCTCAAAATGGAATTTCTGAATATCTGATATCTTCCCTGGACCA	960
Db	1013	AATGTAAATGATCTTCTCAAAATGGAATTTCTGAATATCTGATATCTTCCCTGGACCA	1072
OY	961	GTAGAAATTTTAAACACAGTACAGAAATCAGTTACAGGATATTTCAAGAAATGTAACT	1020
Db	1073	GTAGAAATTTTAAACACAGTACAGAAATCAGTTACAGGATATTTCAAGAAATGTAACT	1132
OY	1021	AATGTGTTCTTCAAGCAACATGATTAACATCATGTAGTTTTTGTATTAATCACTGTCTTG	1080
Db	1133	AATGTGTTCTTCAAGCAACATGATTAACATCATGTAGTTTTTGTATTAATCACTGTCTTG	1192

QY	1081	AAGCATTCTCAACTTTCACAGCCACGTGAAGTAACTGTGTGATGCGCAACCACTG	1140
Db	1193	AAGCATTCTCAACTTTCACAGCCACGTGAAGTAACTGTGTGATGCGCAACCACTG	1252
QY	1141	ACTCTTGAAGAAAGAAATAATGTAGTATAACAGTGAACAAGAAACTATACGTAG	1200
Db	1253	ACTCTTGAAGAAAGAAATAATGTAGTATAACAGTGAACAAGAAACTATACGTAG	1312
QY	1201	TACTGAGCGGATCTAAACGTGAAATCAGAAATGGAAGCTGTTCAAAATTAATTAT	1260
Db	1313	TACTGAGCGGATCTAAACGTGAAATCAGAAATGGAAGCTGTTCAAAATTAATTAT	1372
QY	1261	ACTGTCCCCCAAGTGAACCTTTTAAGATTGAATTCCTCGAGAGATCCAGTAG	1320
Db	1373	ACTGTCCCCCAAGTGAACCTTTTAAGATTGAATTCCTCGAGAGATCCAGTAG	1432
QY	1321	CTACAGTTGAAGGCTATTTCTTGATGATTAAGTAAGTAGACGTGCAGTTCATGCTGTT	1380
Db	1433	CTACAGTTGAAGGCTATTTCTTGATGATTAAGTAAGTAGACGTGCAGTTCATGCTGTT	1492
QY	1381	AAGTCTCTAGTAAAGACATACATCCAACTAATAAACAAGATGAATAATTAAGTGGGA	1440
Db	1493	AAGTCTCTAGTAAAGACATACATCCAACTAATAAACAAGATGAATAATTAAGTGGGA	1552
QY	1441	TGCGCTTTTGAGTTGGTGGTGTAGTGAACAACAAGATTGAAGAGTTAAGCTATATGCTA	1500
Db	1553	TGCGCTTTTGAGTTGGTGGTGTAGTGAACAACAAGATTGAAGAGTTAAGCTATATGCTA	1612
QY	1501	GTATCCAGGGGACAGTTGGTGGCTGTAGGAAACAAATTCACAAATGTTCTCTTTAACA	1560
Db	1613	GTATCCAGGGGACAGTTGGTGGCTGTAGGAAACAAATTCACAAATGTTCTCTTTAACA	1672
QY	1561	CCAGAAATTCCTTGACATCCAAAGCCTGTGTATATGCTATATATTTGAAGATGATGGG	1620
Db	1673	CCAGAAATTCCTTGACATCCAAAGCCTGTGTATATGCTATATATTTGAAGATGATGGG	1732
QY	1621	GAAATTTATTAAGTATGTTCTTAATAATTCCTGTTCAAGCTTGTTTTTAAATAAGATTAAG	1680
Db	1733	GAAATTTATTAAGTATGTTCTTAATAATTCCTGTTCAAGCTTGTTTTTAAATAAGATTAAG	1792
QY	1681	CTATATTTGGAGTAAGTGAAGACGTGAACATCTGAGAAAGTCTCTTAGAGTCTCTGTG	1740
Db	1793	CTATATTTGGAGTAAGTGAAGACGTGAACATCTGAGAAAGTCTCTTAGAGTCTCTGTG	1852
QY	1741	ACACAGCCTGACTCATATGTTGGATTTGATGCTGTGACAAAAGTGTGAATTCGATCAAT	1800
Db	1853	ACACAGCCTGACTCATATGTTGGATTTGATGCTGTGACAAAAGTGTGAATTCGATCAAT	1912
QY	1801	GCCTCTAATGATTTAACAATGGAATAATGTGCTCATGATTTGGAACCTTTAACAACAGA	1860
Db	1913	GCCTCTAATGATTTAACAATGGAATAATGTGCTCATGATTTGGAACCTTTAACAACAGA	1972
QY	1861	TATTTATTTAGGCAATGTTCAATGAATCTTTTGCACTCTTTCAGGAAATGGAACCTCTGGGTA	1920
Db	1973	TATTTATTTAGGCAATGTTCAATGAATCTTTTGCACTCTTTCAGGAAATGGAACCTCTGGGTA	2032
QY	1921	TTGACAGATGAAACCTCAAGAGAGTTATATTGATGCTGTTTATATGCAATGCAAGATAT	1980
Db	2033	TTGACAGATGAAACCTCAAGAGAGTTATATTGATGCTGTTTATATGCAATGCAAGATAT	2092
QY	1981	GCTGAGAGGTTTATGAGAGAAATGAAGGACATATTTGTAGATATTCATGACTTTTCTTGG	2040
Db	2093	GCTGAGAGGTTTATGAGAGAAATGAAGGACATATTTGTAGATATTCATGACTTTTCTTGG	2152
QY	2041	GGTAGCAGTCCACATGTCGGAAGCATTTTCCAGAGACTTGGATTGGCTAGACCAAC	2100
Db	2153	GGTAGCAGTCCACATGTCGGAAGCATTTTCCAGAGACTTGGATTGGCTAGACCAAC	2212
QY	2101	ATGGGTTTACAGGATTTTACAAAGATTTGAAGTAACTGTACCTGATTTATACACTTCTTGG	2160
Db	2213	ATGGGTTTACAGGATTTTACAAAGATTTGAAGTAACTGTACCTGATTTATACACTTCTTGG	2272
QY	2161	GTGCTCTACTGGTTTGTGATCTCTGAGACCTGGGTCCTTGACATTAACACTACTCCAGTG	2220



Db 2273 GTGGCTACGCTGTTTGTGATCTCTGAGACCTGGGGCTTGGACTAACACTACTCCAGAG 2332  
QY 2221 GAGCTCCAGACCTTCCAACTATTTTCATTTTTTGAATCTTCCCTACCTGTATATCAGA 2280  
Db 2333 GAGCTCCAGACCTTCCAACTATTTTCATTTTTTGAATCTTCCCTACCTGTATATCAGA 2392  
QY 2281 GGTGAAGAAATTTGCTTTGGAAATTAATATATTTCAATTTATTTGAAGATCCACTGAGGT 2340  
Db 2393 GGTGAAGAAATTTGCTTTGGAAATTAATATTTCAATTTATTTGAAGATCCACTGAGGT 2452  
QY 2341 AAGGTATATTTGAGAAAAGTGAACAATTTGATTTATTTAATGACTTCAAGTGAATTAAT 2400  
Db 2453 AAGGTATATTTGAGAAAAGTGAACAATTTGATTTATTTAATGACTTCAATGAATTAAT 2512  
QY 2401 GCCACAGGCGACAGAGACCTTCTGTGTCCAGAGAGATGGGGCACTGTCTTTT 2460  
Db 2513 GCCACAGGCGACAGAGACCTTCTGTGTCCAGAGAGATGGGGCACTGTCTTTT 2572  
QY 2461 CCCATCAGGCGAACACATCTGGGAGAAATTCCTATCAGTCAAGCTCTTCACCACT 2520  
Db 2573 CCCATCAGGCGAACACATCTGGGAGAAATTCCTATCAGTCAAGCTCTTCACCACT 2632  
QY 2521 GCTTCTGATGCTGTCAACCAAGATTTTATGTAAGGCTGAAGGAATGAAATATCATAT 2580  
Db 2633 GCTTCTGATGCTGTCAACCAAGATTTTATGTAAGGCTGAAGGAATGAAATATCATAT 2692  
QY 2581 TCACATTCATCTTATTTTATGACTGACATAGGCTACAGATACCTGAAAACCTTG 2640  
Db 2693 TCACATTCATCTTATTTTATGACTGACATAGGCTACAGATACCTGAAAACCTTG 2752  
QY 2641 AGTTTCTCATTTCTCTCTAATACAGTGACTGGCAGTGAAGAAAGTTTCAGTCACTGCAAT 2700  
Db 2753 AGTTTCTCATTTCTCTCTAATACAGTGACTGGCAGTGAAGAAAGTTTCAGTCACTGCAAT 2812  
QY 2701 GGAGATGTTCTTGTCTTCCATCAATGAGCTTACCTCATGATGAGATGAGCTTAATGCG 2760  
Db 2813 GGAGATGTTCTTGTCTTCCATCAATGAGCTTACCTCATGATGAGATGAGCTTAATGCG 2872  
QY 2761 TGTGTGAACAGAACATGATAAATTTTGTCCAAAATTTTACATTTTGAATATCTGACT 2820  
Db 2873 TGTGTGAACAGAACATGATAAATTTTGTCCAAAATTTTACATTTTGAATATCTGACT 2932  
QY 2821 AAAAAAGAAACAATGACAGATTAATTTGAAGAAAAGCTCTTTCATTTATGAGGCAAGGT 2880  
Db 2933 AAAAAAGAAACAATGACAGATTAATTTGAAGAAAAGCTCTTTCATTTATGAGGCAAGGT 2992  
QY 2881 TACAGAGAGAACTTCTCTATCAGAGAGAGATGAGCTCTTTCAGTGCCTTTTGGAAATTAAT 2940  
Db 2993 TACAGAGAGAACTTCTCTATCAGAGAGAGATGAGCTCTTTCAGTGCCTTTTGGAAATTAAT 3052  
QY 2941 GACCCCTTCTGGAGACCTTGGTGTGACCTTTTGAAGATGTTTCTTGAAGCCGAT 3000  
Db 3053 GACCCCTTCTGGAGACCTTGGTGTGACCTTTTGAAGATGTTTCTTGAAGCCGAT 3112  
QY 3001 CCTTACATGATATTTGATCAGATGATGTTTACACAGAACATACCTTGCTTAAGAGCAAT 3060  
Db 3113 CCTTACATGATATTTGATCAGATGATGTTTACACAGAACATACCTTGCTTAAGAGCAAT 3172  
QY 3061 CAGAAATCCAAAGGTGAATTTTGGATCCAGAGAAAGATTCATAGTAGCTTCAAGGT 3120  
Db 3173 CAGAAATCCAAAGGTGAATTTTGGATCCAGAGAAAGATTCATAGTAGCTTCAAGGT 3232  
QY 3121 GGCATTAAGATTCAGTAACTTACAGCTTATTTGATCTTCTCTGGAGATTAAGA 3180  
Db 3233 GGCATTAAGATTCAGTAACTTACAGCTTATTTGATCTTCTCTGGAGATTAAGA 3292  
QY 3181 AAGTATCAGCTTAACATGATGCAAGAGTCAATCTTTTGGAGCTGAATTCAGT 3240  
Db 3293 AAGTATCAGCTTAACATGATGCAAGAGTCAATCTTTTGGAGCTGAATTCAGT 3352  
QY 3241 AAGAGAAATTCAGACAATTAATCTAGCCCTTAATACCTTATGATGCTACAGTGGGG 3300

Db 3353 AAGAGAAATTCAGACAATTAATCTTAGCCCTTAATTAATGATGCTACAGTGGGG 3412  
QY 3301 AGTCTTAAGAGAGAGAGAGCTTTGAAATATGCTGACTTGAAGAGAGAACAAAGAGTGGC 3360  
Db 3413 AGTCTTAAGAGAGAGAGAGCTTTGAAATATGCTGACTTGAAGAGAGAACAAAGAGTGGC 3472  
QY 3361 ATGCAATTCGGGTGTCAACAGAGTCAAACTTTCTGACTCCTGGCAGCAGCTCCCTG 3420  
Db 3473 ATGCAATTCGGGTGTCAACAGAGTCAAACTTTCTGACTCCTGGCAGCAGCTCCCTG 3532  
QY 3421 GATATTAAGATTCAGCTTATGCACTGCTCTCAACTTTCTTAACAATTTCAACTTTGAG 3480  
Db 3533 GATATTAAGATTCAGCTTATGCACTGCTCTCAACTTTCTTAACAATTTCAACTTTGAG 3592  
QY 3481 GGAATCCCAATTAAGAGTGGCTTAAGCAGAGAAAGAAATACCTGGGTGGTTTGCACT 3540  
Db 3593 GGAATCCCAATTAAGAGTGGCTTAAGCAGAGAAAGAAATACCTGGGTGGTTTGCACT 3652  
QY 3541 ACTCAGATTAACCACTGTGGCTTTAAAGGCTGTCTGAAATTTGAGCCCTTAATGAATACA 3600  
Db 3653 ACTCAGATTAACCACTGTGGCTTTAAAGGCTGTCTGAAATTTGAGCCCTTAATGAATACA 3712  
QY 3601 GAAAGCAAAATATTCAAAGTACCGGTGACGGGGCTTAGCTCAACCAAGTCT----- 3651  
Db 3713 GAAAGCAAAATATTCAAAGTACCGGTGACGGGGCTTAGCTCAACCAAGTCTGTAAGTTT 3772  
QY 3652 -----CTTGTGTGTGACGCA 3669  
Db 3773 CTGATTAACAACAACAACCGCTTACTCTTCAAGACAGAGAGCTTGTGTGACGCA 3832  
QY 3670 ATGCAAGTTAATTAATTTCCGAAATGTTTGAATTTGATTTGTCAGCTCAATGTTGTA 3729  
Db 3883 ATGCAAGTTAATTAATTTCCGAAATGTTTGAATTTGTCAGCTCAATGTTGTA 3892  
QY 3730 TATAATGTAAGGCTTGTGGGTCTTTCAGAGACGAGAGATCTATCCAAATCAAGAACCC 3789  
Db 3893 TATAATGTAAGGCTTGTGGGTCTTTCAGAGACGAGAGATCTATCCAAATCAAGAACCC 3952  
QY 3790 TTTGATTTAAGATGTTGCTGTAAAGAAATTAAGATGATCTCAATCTAGTGTGAAT 3849  
Db 3953 TTTGATTTAAGATGTTGCTGTAAAGAAATTAAGATGATCTCAATCTAGTGTGAAT 4012  
QY 3850 GTGTGTACAAGCTTTTGGGGCCGGGTGAGAGTGCATGAGCTTATGGAAGTTAACCTA 3909  
Db 4013 GTGTGTACAAGCTTTTGGGGCCGGGTGAGAGTGCATGAGCTTATGGAAGTTAACCTA 4072  
QY 3910 TTAAGTGCCTTATGAGTCTTTCAGAGCAATTTTCTGACGAGACAGTGAAGAAAGT 3969  
Db 4073 TTAAGTGCCTTATGAGTCTTTCAGAGCAATTTTCTGACGAGACAGTGAAGAAAGT 4132  
QY 3970 GAATATGATTCATGAGAAAATCAACTCTATTTAATCTGTAAATGAAAACCAAGTTTGT 4029  
Db 4133 GAATATGATTCATGAGAAAATCAACTCTATTTAATCTGTAAATGAAAACCAAGTTTGT 4192  
QY 4030 GTTAATTAATTCGTGCTGAGAGAACTTTAAAGTTTCAATFACCAAGATGCTTCAAGTCC 4089  
Db 4193 GTTAATTAATTCGTGCTGAGAGAACTTTAAAGTTTCAATFACCAAGATGCTTCAAGTCC 4252  
QY 4090 ATAGTGAATTAATTAAGACCAAGAGACAGCGGTGAGAAATTAACAATCTGAAGTGAAG 4149  
Db 4253 ATAGTGAATTAATTAAGACCAAGAGACAGCGGTGAGAAATTAACAATCTGAAGTGAAG 4312  
QY 4150 CTGTCTCTCTGTGACCTTTGCAAGTATGTCAGAGGCTGCGCTTGTGTGAGAGTGAAGT 4209  
Db 4313 CTGTCTCTCTGTGACCTTTGCAAGTATGTCAGAGGCTGCGCTTGTGTGAGAGTGAAGT 4372  
QY 4210 TCAGGCTCCCAATCACTCTTCAAGTATTTTATTTTCTGTTCAGAGCTTCTGACTTT 4269  
Db 4373 TCAGGCTCCCAATCACTCTTCAAGTATTTTATTTTCTGTTCAGAGCTTCTGACTTT 4432  
QY 4270 ATGGAACCTTGGCTGTGA 4287  
Db 4433 ATGGAACCTTGGCTGTGA 4450

RESULT 6  
 ABQ79965  
 ID ABQ79965 standard; cDNA: 4761 BP.  
 XX AC ABQ79965;  
 XX AC ABQ79965;  
 DT 23-DEC-2002 (first entry)  
 XX  
 DE Human CD109 K1 variant protein encoding cDNA.  
 XX  
 KM CD109; alpha2 macroglobulin; thioester; cerebroprotective; cardiact;  
 KM immunosuppressive; haemostatic; anticoagulant; thrombolytic; human;  
 KM cardiovascular; vasotropic; gene therapy; CD109 K1; variant; gene; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 FT CDS 113..4450  
 FT /\*tag= a  
 FT /product= "CD109 K1 variant"  
 XX  
 XX MO200270696-A2.  
 XX  
 XX 12-SEP-2002.  
 XX  
 XX 07-MAR-2002; 2002MO-CA00292.  
 XX  
 XX 07-MAR-2001; 2001US-273814P.  
 XX  
 XX (SCHU/) SCHUH A.  
 XX (SUTH/) SUTHERLAND R D.  
 XX  
 XX Schuh A, Sutherland RD;  
 XX  
 XX WPI; 2002-713450/77.  
 DR P-PSDB; ABB82166.  
 XX  
 XX  
 PT New CD109 nucleic acids and polypeptides, useful in gene therapy,  
 PT particularly for treating strokes, myocardial infarctions, thrombosis,  
 PT thrombocytopenia, autoimmune diseases, or organ or bone marrow  
 PT transplantation -  
 XX  
 XX  
 PS Claim 1; Fig 1b; 156bp; English.  
 XX  
 XX The invention relates to isolated nucleic acid molecules encoding CD109  
 XX polypeptides. These nucleic acid molecules include the human cDNA  
 XX sequences comprising CD109 K1, CD109 K1-H7, CD109 K15 or their variants.  
 XX CD109 is a novel member of the alpha2 macroglobulin (alpha2M)/C3, C4, C5  
 XX family of thioester-containing proteins. The CD109 polypeptides can be  
 XX expressed by standard recombinant methodology. The CD109 nucleic acid,  
 XX CD109 polypeptide or its mimetic, or the anti-CD109 antibody is useful  
 XX for treating or detecting a disease or disorder, e.g. conditions  
 XX associated with endothelial activation, platelet activation, activation  
 XX and of the coagulation or fibrinolytic systems, activation of T lymphocytes  
 XX and of the complement system, quantitative or qualitative abnormalities  
 XX of platelet function, increased or impaired platelet aggregation and  
 XX activation, increased or impaired activation of the coagulation and/or  
 XX fibrinolytic systems, or impaired or increased immune activation. These  
 XX are also useful for treating cardiovascular disorders, stroke, myocardial  
 XX infarction, thrombosis, embolism, peripheral vascular disease, organ  
 XX thrombocytopenia, thrombocythemia, autoimmune diseases, organ  
 XX transplantation, or bone marrow transplantation. The CD109 nucleic acids  
 XX or their antisense nucleotide sequence are useful in gene therapy for  
 XX treating these conditions. The present sequence represents the human  
 XX CD109 K1 variant cDNA sequence.  
 XX  
 SQ Sequence 4761 BP; 1419 A; 913 C; 1016 G; 1413 T; 0 other;  
 Query Match 98.5%; Score 4222.8; DB 24; Length 4761;  
 Best Local Similarity 98.8%; Pred. No. 0;  
 Matches 4285; Conservative 0; Mismatches 2; Indels 51; Gaps 1;

Qy	1	ATCAGAGGCCACCGCTCTGACCGCCGCCACCTCTCTGCGTGTGACCGCCGCGCTG	60
Db	113	ATCAGAGGCCACCGCTCTGACCGCCGCCACCTCTCTGCGTGTGACCGCCGCGCTG	172
Qy	61	GCCGTGGCTCCCGGCGCTGGCTTCTGTGTGACAGCCCGGAGATCATACAGCCCGGAGGA	120
Db	173	GCCGTGGCTCCCGGCGCTGGCTTCTGTGTGACAGCCCGGAGATCATACAGCCCGGAGGA	232
Qy	121	AATGTAGTATTGGGGGAGGCTTGTGAACACTGCCCTTACAGGTGACTGTGAAGGCG	180
Db	233	AATGTAGTATTGGGGGAGGCTTGTGAACACTGCCCTTACAGGTGACTGTGAAGGCG	292
Qy	181	GACCTGCTCAAGACAGATCAAACTCACTGTCTGTCTGTGAGACAGAGAGTCTTT	240
Db	293	GACCTGCTCAAGACAGATCAAACTCACTGTCTGTCTGTGAGACAGAGAGTCTTT	352
Qy	241	GAAGAAAGGCTCTTTTAAGACATTACTCTTCCATCTACCTTGAACAGTGAATGAG	300
Db	353	GAAGAAAGGCTCTTTTAAGACATTACTCTTCCATCTACCTTGAACAGTGAATGAG	412
Qy	301	ATTATAGCTAGCTGTAAACCGGAGCTACCCAGATGAGATTTATCTTAATAGTACC	360
Db	413	ATTATAGCTAGCTGTAAACCGGAGCTACCCAGATGAGATTTATCTTAATAGTACC	472
Qy	361	CGCTTATCATTTTGAGACCAAGAGATATCTGTCTTCAATCAACAGACAGGCTTATAC	420
Db	473	CGCTTATCATTTTGAGACCAAGAGATATCTGTCTTCAATCAACAGACAGGCTTATAC	532
Qy	421	AAGCCAAAGCAGAGATGAACTTGGCATTTGTACCTCTTCAATTTTAAGCTTATC	480
Db	533	AAGCCAAAGCAGAGATGAACTTGGCATTTGTACCTCTTCAATTTTAAGCTTATC	592
Qy	481	AAACCTCTTTAAACATCTCATTAAGAGACCCCAATCAATTTGATCCACAGTGTG	540
Db	593	AAACCTCTTTAAACATCTCATTAAGAGACCCCAATCAATTTGATCCACAGTGTG	652
Qy	541	TCACAAACAAAGTATCTTGAGATCATTTCCAAACTTTCAAGTATCTTCCATCCATTA	600
Db	653	TCACAAACAAAGTATCTTGAGATCATTTCCAAACTTTCAAGTATCTTCCATCCATTA	712
Qy	601	CTTGGTACTGCTATTTCAAGTTCAAGTATGACCAAGATATTAATCATTTTCAG	660
Db	713	CTTGGTACTGCTATTTCAAGTTCAAGTATGACCAAGATATTAATCATTTTCAG	772
Qy	661	GTTTCAATATGTATTAACCAAAATTGAAGTACTTGCAGACACCATTAATGTCT	720
Db	773	GTTTCAATATGTATTAACCAAAATTGAAGTACTTGCAGACACCATTAATGTCT	832
Qy	721	ATGAATTTCAAGCATTTAATGTGTACATCAGCGCAAGATATCATATGGAAGCCAGTG	780
Db	833	ATGAATTTCAAGCATTTAATGTGTACATCAGCGCAAGATATCATATGGAAGCCAGTG	892
Qy	781	AAAGAGAGGTAAAGCTTACATTTTACCTTATCTTTGGGAAAAGAAAATATT	840
Db	893	AAAGAGAGGTAAAGCTTACATTTTACCTTATCTTTGGGAAAAGAAAATATT	952
Qy	841	ACAAAAACATTAAGTAAATGSAATCTGCAACTCTTTTAATGATGAAGATGA	900
Db	953	ACAAAAACATTAAGTAAATGSAATCTGCAACTCTTTTAATGATGAAGATGA	1012
Qy	901	AATGTATGATTTCTCAATGATGACTTGTGAATACCTGATCTATCTTCCCTGAGCA	960
Db	1013	AATGTATGATTTCTCAATGATGACTTGTGAATACCTGATCTATCTTCCCTGAGCA	1072
Qy	961	GTAGAAATTTTAAACACAGTACAGAAATGATTAAGTATTTCAAGAAATGATACACT	1020
Db	1073	GTAGAAATTTTAAACACAGTACAGAAATGATTAAGTATTTCAAGAAATGATACACT	1132
Qy	1021	AATGTGTCTTCAAGCAATGATTAATCATTAATGATTTTGAATATATCTGTCTG	1080
Db	1133	AATGTGTCTTCAAGCAATGATTAATCATTAATGATTTTGAATATATCTGTCTG	1192

QY 1081 AAGCCATCTCTCACTTCAAGCACTGTGAAAGTAACTGCTGATGATGCAACCACTG 1140  
DB 1193 AAGCCATCTCTCACTTCAAGCACTGTGAAAGTAACTGCTGATGATGCAACCACTG 1252  
QY 1141 AACTGTGAAGAAAGAAATAATGATGATCACTAAGTGAACAGAGAACTATCTGAG 1200  
DB 1253 ACTCTTGAAGAAAGAAATAATGATGATCACTAAGTGAACAGAGAACTATCTGAG 1312  
QY 1201 TACTGAGCGGATCTAAGCAGTGAATCAGAAAATGGAAGCTGTCAGAAAATTAATAT 1260  
DB 1313 TACTGAGCGGATCTAAGCAGTGAATCAGAAAATGGAAGCTGTCAGAAAATTAATAT 1372  
QY 1261 ACTGTCCCCCAAGTGAACCTTTAAGATGAATCCCAATCTGAGAGATTCAGTGA 1320  
DB 1373 ACTGTCCCCCAAGTGAACCTTTAAGATGAATCCCAATCTGAGAGATTCAGTGA 1432  
QY 1321 CTACAGTTGAAGGCCATTTCTGCTGATGAATAAGATGAGAGATTCATAGTCTGTT 1380  
DB 1433 CTACAGTTGAAGGCCATTTCTGCTGATGAATAAGATGAGAGATTCATAGTCTGTT 1492  
QY 1381 AAGTCTCTAGTAAAGATACATCCAACTAAAAACAAGATGAATAATTAAGTGA 1440  
DB 1493 AAGTCTCTAGTAAAGATACATCCAACTAAAAACAAGATGAATAATTAAGTGA 1552  
QY 1441 TGGCTTTTGAAGTGTGATGATGAGCAACAAGATGAAGATTAAGTGA 1500  
DB 1553 TGGCTTTTGAAGTGTGATGATGAGCAACAAGATGAAGATTAAGTGA 1612  
QY 1501 GATTCAGAGGACAGTGTGCTGATGAGAAACAAATTCACAACTGCTCTTTACA 1560  
DB 1613 GATTCAGAGGACAGTGTGCTGATGAGAAACAAATTCACAACTGCTCTTTACA 1672  
QY 1561 CCAGAAAATTTCTGATCTCAAGCCGTGTGATGATTAATGAAGATGATG 1620  
DB 1673 CCAGAAAATTTCTGATCTCAAGCCGTGTGATGATTAATGAAGATGATG 1732  
QY 1621 GAATTAATGAAGTGTCTTAATAAATCTGCTCACTGTTTAAATTAAGTGA 1680  
DB 1733 GAATTAATGAAGTGTCTTAATAAATCTGCTCACTGTTTAAATTAAGTGA 1792  
QY 1681 CTATATGAGTAAAGTGAAGTGAACCACTGAGAAAGTCTCTTAAAGTCTG 1740  
DB 1793 CTATATGAGTAAAGTGAAGTGAACCACTGAGAAAGTCTCTTAAAGTCTG 1852  
QY 1741 ACACAGCCGATCCATAGTGGATGATGATGATGATGATGATGATGATGAT 1800  
DB 1853 ACACAGCCGATCCATAGTGGATGATGATGATGATGATGATGATGATGAT 1912  
QY 1801 GCTCTAATGATTAACAAATGAAGTGTGATGATGATGATGATGATGATGAT 1860  
DB 1913 GCTCTAATGATTAACAAATGAAGTGTGATGATGATGATGATGATGATGAT 1972  
QY 1861 TATTAATGAGCATGATGATGATGATGATGATGATGATGATGATGATGAT 1920  
DB 1973 TATTAATGAGCATGATGATGATGATGATGATGATGATGATGATGATGAT 2032  
QY 1921 TTGACAGATGCAAACTCAAGAAATTAATGATGATGATGATGATGATGAT 1980  
DB 2033 TTGACAGATGCAAACTCAAGAAATTAATGATGATGATGATGATGATGAT 2092  
QY 1981 GCTGAGAGGTTATGAGAGAAATGAAGACATATGATGATGATGATGATGAT 2040  
DB 2093 GCTGAGAGGTTATGAGAGAAATGAAGACATATGATGATGATGATGATGAT 2152  
QY 2041 GGTAGAGTCCACATGTCGAAAGATTTTCAAGAGATGATGATGATGATGAT 2100  
DB 2153 GGTAGAGTCCACATGTCGAAAGATTTTCAAGAGATGATGATGATGATGAT 2212  
QY 2101 ATGGGTTAAGATTAACAAAGATTTGAAGTGAACGATGATGATGATGATGAT 2160  
DB 2213 ATGGGTTAAGATTAACAAAGATTTGAAGTGAACGATGATGATGATGATGAT 2272  
QY 2161 GTGGCTACTGTTTGTGATCTCTGAGACCTGGGTCTTGAACACTACTCAGTG 2220

DB 2273 GTGGCTACTGTTTGTGATCTCTGAGACCTGGGTCTTGAACACTACTCAGTG 2332  
QY 2221 GAGCTCCAAAGCTTCCAAACATTTTCAATTTTGAATCTTCCCTACTGATTAAGA 2280  
DB 2333 GAGCTCCAAAGCTTCCAAACATTTTCAATTTTGAATCTTCCCTACTGATTAAGA 2392  
QY 2281 GGTGAAGATTTGCTTGAAGATTAATGATTAATTAATTAATTAATTAATTAAT 2340  
DB 2393 GGTGAAGATTTGCTTGAAGATTAATGATTAATTAATTAATTAATTAATTAAT 2452  
QY 2341 AAGTAAATCATTGAGAAAGTGAACAAATTTGATTAATGATGATGATGATGAT 2400  
DB 2453 AAGTAAATCATTGAGAAAGTGAACAAATTTGATTAATGATGATGATGATGAT 2512  
QY 2401 GCCACAGGCCACAGAGACCTTCTGCTCCAGTGAAGATGAGAGGCACTGTTT 2460  
DB 2513 GCCACAGGCCACAGAGACCTTCTGCTCCAGTGAAGATGAGAGGCACTGTTT 2572  
QY 2461 CCCATCAGGCCACACATCTGGAGAAATTCATCAAGTCAAGCTCTTCAACCACT 2520  
DB 2573 CCCATCAGGCCACACATCTGGAGAAATTCATCAAGTCAAGCTCTTCAACCACT 2632  
QY 2521 GCTTGTGATGCTGACCCAGATGATTTAGTAAAGCTGAAGATGAAGAAATCATAT 2580  
DB 2633 GCTTGTGATGCTGACCCAGATGATTTAGTAAAGCTGAAGATGAAGAAATCATAT 2692  
QY 2581 TCACATCAATCTTATTAATGATGATGATGATGATGATGATGATGATGATGAT 2640  
DB 2693 TCACATCAATCTTATTAATGATGATGATGATGATGATGATGATGATGATGAT 2752  
QY 2641 AGTTCTCATTTCTCCCTAATAGATGATGATGATGATGATGATGATGATGAT 2700  
DB 2753 AGTTCTCATTTCTCCCTAATAGATGATGATGATGATGATGATGATGATGAT 2812  
QY 2701 GAGATGTTCTTGTGCTTCCATCAATGATGATGATGATGATGATGATGATGAT 2760  
DB 2813 GAGATGTTCTTGTGCTTCCATCAATGATGATGATGATGATGATGATGATGAT 2872  
QY 2761 TGTGTGAACGAACATGATTAATTTGCTCAAAATTTACATTTGATGATGATGAT 2820  
DB 2873 TGTGTGAACGAACATGATTAATTTGCTCAAAATTTACATTTGATGATGATGAT 2932  
QY 2821 AAAAAAACAACGACGATTAATTTGAAGAAAGCTCTTCAATTAAGGCAAGT 2880  
DB 2933 AAAAAAACAACGACGATTAATTTGAAGAAAGCTCTTCAATTAAGGCAAGT 2992  
QY 2881 TACAGAGAGAACTTCTCTATCAGAGGAGATGCTCTTCAAGTCTTTGGGAATTA 2940  
DB 2993 TACAGAGAGAACTTCTCTATCAGAGGAGATGCTCTTCAAGTCTTTGGGAATTA 3052  
QY 2941 GACCTTCTGGAGACCTGTTGTCACTTTGTTTAAAGTGTTCCTTGAACCGAT 3000  
DB 3053 GACCTTCTGGAGACCTGTTGTCACTTTGTTTAAAGTGTTCCTTGAACCGAT 3112  
QY 3001 CCTTACATGATTAATGATCAAGATGATTAACGAACATCACTGCTTAAAGGACAT 3060  
DB 3113 CCTTACATGATTAATGATCAAGATGATTAACGAACATCACTGCTTAAAGGACAT 3172  
QY 3061 CAGAAATCAACGGGATTTTGGGATCAGAGAGATGATTAAGTGAAGCTTCAAGT 3120  
DB 3173 CAGAAATCAACGGGATTTTGGGATCAGAGAGATGATTAAGTGAAGCTTCAAGT 3232  
QY 3121 GGCATTAATAAGTCCAGTAACTTACAGCTTAATTTGAATCTTCTCTGGATTAAGA 3180  
DB 3233 GGCATTAATAAGTCCAGTAACTTACAGCTTAATTTGAATCTTCTCTGGATTAAGA 3292  
QY 3181 AAGTAAAGCTTAATGATGATGATGATGATGATGATGATGATGATGATGAT 3240  
DB 3293 AAGTAAAGCTTAATGATGATGATGATGATGATGATGATGATGATGATGAT 3352  
QY 3241 AAGAAATTTGAGCAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3300

Db	3353	AGAGAAATTTGACGAATTTATACTGTACCCCTTAAACTTATGCAATGTTCATGAGTGGG	3412
Qy	3301	AGTCTTAAGCGAAGAGCTTTGAAATATGCTGACTTGGAGACGAACAAAGAGTGGC	3360
Db	3413	AGTCTTAAGCGAAGAGCTTTGAAATATGCTGACTTGGAGACGAACAAAGAGTGGC	3472
Qy	3361	ATGCAAATCTGGGTGCATCAGAGTCCAAACCTTCTGACTCTGGACGACGCTCCCTG	3420
Db	3473	ATGCAAATCTGGGTGCATCAGAGTCCAAACCTTCTGACTCTGGACGACGCTCCCTG	3532
Qy	3421	GATATTGAGTTGACGCTATGCACTGCTCTCACACTTCTTACAATTCAGACTTGTAG	3480
Db	3533	GATATTGAGTTGACGCTATGCACTGCTCTCACACTTCTTACAATTCAGACTTGTAG	3592
Qy	3481	GGAATCCCAATTAATGAGGTGGCTAAGCAGCGCAAGAAATACCTGGGTGTTTGCATCT	3540
Db	3593	GGAATCCCAATTAATGAGGTGGCTAAGCAGCGCAAGAAATACCTGGGTGTTTGCATCT	3652
Qy	3541	ACTCAGATACCACTGCTGGCTTTAAAGGCTGCTGATGAAATTTGGACGCCCTAATGATACA	3600
Db	3653	ACTCAGATACCACTGCTGGCTTTAAAGGCTGCTGATGAAATTTGGACGCCCTAATGATACA	3712
Qy	3601	GAAAGACAATATTCAGATGACCGCTGACGGGGCTAGCTACCAAGTCTC-----	3651
Db	3713	GAAAGACAATATTCAGATGACCGCTGACGGGGCTAGCTACCAAGTCTC-----	3772
Qy	3652	-----CTTGCTGTGTACAGCCA	3669
Db	3773	CTGATTTGACACACACACACCGCTTACTCCTTCAGACAGCAGAGCTGCTGTGTACAGCCA	3832
Qy	3670	ATGGCAGTTAATATTTCCGCAAAAGGTTTGGATTGCTATTTGTCAGCTCAAGTTGTA	3729
Db	3833	ATGGCAGTTAATATTTCCGCAAAAGGTTTGGATTGCTATTTGTCAGCTCAAGTTGTA	3892
Qy	3730	TATATATGGAAGGCTTCTGGTCTTCTTGAAGAAGAGATCTATCCAAATCAAGAAACC	3789
Db	3893	TATATATGGAAGGCTTCTGGTCTTCTTGAAGAAGAGATCTATCCAAATCAAGAAACC	3952
Qy	3790	TTTGAATTTAGATGTTGCTGTAAAGAAATAAAGATGATCTCAATCATGTGGAATTTGAT	3849
Db	3953	TTTGAATTTAGATGTTGCTGTAAAGAAATAAAGATGATCTCAATCATGTGGAATTTGAT	4012
Qy	3850	GTGTGTACAACTTTTCCGGCCCGGGTATGAGTGCATGCGCTTTATGGAATTTACCTTA	3909
Db	4013	GTGTGTACAACTTTTCCGGCCCGGGTATGAGTGCATGCGCTTTATGGAATTTACCTTA	4072
Qy	3910	TTAAGTGGCTTTATGAGTGCCTTTCAGAAACAATTTCTGTGACGAGACAAGTAAAGAAAGTG	3969
Db	4073	TTAAGTGGCTTTATGAGTGCCTTTCAGAAACAATTTCTGTGACGAGACAAGTAAAGAAAGTG	4132
Qy	3970	GAATATGATCATGGAATACTCAACCTCTATTTAGATTCTGTAAATGAAACCAGTTTGT	4029
Db	4133	GAATATGATCATGGAATACTCAACCTCTATTTAGATTCTGTAAATGAAACCAGTTTGT	4192
Qy	4030	GTTTAATATTTCTGCTGTGTGAGAAACTTTTAAAGTTTCAATTAACCAAGATGCTTCAGTGTCC	4089
Db	4193	GTTTAATATTTCTGCTGTGTGAGAAACTTTTAAAGTTTCAATTAACCAAGATGCTTCAGTGTCC	4252
Qy	4090	ATTAGTGATTACTATGAGCCAAAGGAGACAGCGCGGTGAGAAAGTTTCAACTCTGAAGTGAAG	4149
Db	4253	ATTAGTGATTACTATGAGCCAAAGGAGACAGCGCGGTGAGAAAGTTTCAACTCTGAAGTGAAG	4312
Qy	4150	CTGTCTCTCTGTGACCTTTGACAGTATGTCAGAGGCTGCGCTCTTGTGAGAGATGAGACT	4209
Db	4313	CTGTCTCTCTGTGACCTTTGACAGTATGTCAGAGGCTGCGCTCTTGTGAGAGATGAGACT	4372
Qy	4210	TCAGGCTTCCATCATCACTCTTCAGTCAATTTTATTTTCTGTTCAAGCTTCTGTACTTT	4269
Db	4373	TCAGGCTTCCATCATCACTCTTCAGTCAATTTTATTTTCTGTTCAAGCTTCTGTACTTT	4432
Qy	4270	ATGGAACCTTTGGCTGTGA	4287
Db	4433	ATGGAACCTTTGGCTGTGA	4450

Query March	Best Local Similarity	98.5%	Score 4222.8	DB 24	Length 5895
Matches 4285	Conservative 0	Mismatches 2	Indels 51	Gaps 1	

QY 1 ATGAGGGCCACCGCTCTGACCGCGCCACCTCTCTGCGTGTGACCGCGCGCTG 60  
DB 113 ATGAGGGCCACCGCTCTGACCGCGCCACCTCTCTGCGTGTGACCGCGCGCTG 172  
QY 61 GCGGTGGCTCCCGGGCTCTGCTTCTGTGTGACCGCGCCAGGATCATGCGCGCGAGA 120  
DB 173 GCGGTGGCTCCCGGGCTCTGCTTCTGTGTGACCGCGCCAGGATCATGCGCGCGAGA 232  
QY 121 AATGTGACTATTGGGGTGTGAGCTTCTGTGAACAGTGGCCCTTCAAGGTGACTGTGAAGGG 180  
DB 223 AATGTGACTATTGGGGTGTGAGCTTCTGTGAACAGTGGCCCTTCAAGGTGACTGTGAAGGG 292  
QY 181 GAGCTGCTCAAGACAGATCAAACTCACTGTCTCTGTCTGTGAAGCAAGAGAGCTTT 240  
DB 293 GAGCTGCTCAAGACAGATCAAACTCACTGTCTCTGTCTGTGAAGCAAGAGAGCTTT 352  
QY 241 GAAAAAGGCTCTTTTAAGACACTTACTCTTCATCACTACCTCTGTGAAGCAAGAGAG 300  
DB 353 GAAAAAGGCTCTTTTAAGACACTTACTCTTCATCACTACCTCTGTGAAGCAAGAGAG 412  
QY 301 ATTATGAGCTAGGTGAACCGGAGTACCGAGAGTAACTTTATCTCTAATAGTACC 360  
DB 413 ATTATGAGCTAGGTGAACCGGAGTACCGAGAGTAACTTTATCTCTAATAGTACC 472  
QY 361 CGCTTATCATTTGAGACCAAGAGAAATCTGTCTTCAATCAACAGACAGAGGCTTATAC 420  
DB 473 CGCTTATCATTTGAGACCAAGAGAAATCTGTCTTCAATCAACAGACAGAGGCTTATAC 532  
QY 421 AAGCCAAAGCAAGAGTGAAGTTGGCATTTGTAACCTCTCTCAAGATTTAAGCCTTAC 480  
DB 533 AAGCCAAAGCAAGAGTGAAGTTGGCATTTGTAACCTCTCTCAAGATTTAAGCCTTAC 592  
QY 481 AAAACCTCTTAAACATTTCTCATTAAGAGCCCAATCAATTTGATCCAAAGAGTGG 540  
DB 593 AAAACCTCTTAAACATTTCTCATTAAGAGCCCAATCAATTTGATCCAAAGAGTGG 652  
QY 541 TCACAAACAAAGTATCTGAGTGCATTTCCAAAACCTTTCAAGTATCTTCCATCAATA 600  
DB 653 TCACAAACAAAGTATCTGAGTGCATTTCCAAAACCTTTCAAGTATCTTCCATCAATA 712  
QY 601 CTGTGTGACTGTGTATTCAGTTCAGTGAATGACCAAGATATTTATCATTTGAG 660  
DB 713 CTGTGTGACTGTGTATTCAGTTCAGTGAATGACCAAGATATTTATCATTTGAG 772  
QY 661 GTTTTGAAGATATGTATTAACAAATTTGAAAGTACTTTGACAGACACCATTAATTTGCT 720  
DB 773 GTTTTGAAGATATGTATTAACAAATTTGAAAGTACTTTGACAGACACCATTAATTTGCT 832  
QY 721 ATGAATTCCTAAGCATTTAAATGTATCAATCAAGGCAAGATATCATTTGAGAGGAGTGG 780  
DB 833 ATGAATTCCTAAGCATTTAAATGTATCAATCAAGGCAAGATATCATTTGAGAGGAGTGG 892  
QY 781 AAAAGAGACGTAAACGCTTACATTTTACCTTTATCTTTTGGGAAAGAGAAAAATAT 840  
DB 893 AAAAGAGACGTAAACGCTTACATTTTACCTTTATCTTTTGGGAAAGAGAAAAATAT 952  
QY 841 ACAAAGACATTTAAATTAATGTATCTGTGAACCTTCTTTTAAAGTAAAGATGA 900  
DB 953 ACAAAGACATTTAAATTAATGTATCTGTGAACCTTCTTTTAAAGTAAAGATGA 1012  
QY 901 AATGTAAATGATTTCTTCAAAATGAGCTTTCTGAATACCTGAGATCTAATCTCCCTGAGACA 960  
DB 1013 AATGTAAATGATTTCTTCAAAATGAGCTTTCTGAATACCTGAGATCTAATCTCCCTGAGACA 1072  
QY 961 GTAGAAATTTTAAACACAGTGAACAGATCAATCAAGTATTTCAAGAAATGTAAAGACT 1020  
DB 1073 GTAGAAATTTTAAACACAGTGAACAGATCAATCAAGTATTTCAAGAAATGTAAAGACT 1132  
QY 1021 AATGTGTCTTCAAGCAACATGATTAATCATTTGAGTTTGTGATTAATCACTGTCTTG 1080  
DB 1133 AATGTGTCTTCAAGCAACATGATTAATCATTTGAGTTTGTGATTAATCACTGTCTTG 1192

QY 1081 AAGCATCTCTCAACTTTCACAGCACTGTGAAGGTAACTGTGTGATGAGCAACCACTG 1140  
DB 1193 AAGCATCTCTCAACTTTCACAGCACTGTGAAGGTAACTGTGTGATGAGCAACCACTG 1252  
QY 1141 ACTTTGAAGAAAGAAATTAATGTAGTCAATAAGTGAACACAGAAACCTATCTGAG 1200  
DB 1253 ACTTTGAAGAAAGAAATTAATGTAGTCAATAAGTGAACACAGAAACCTATCTGAG 1312  
QY 1201 TACTGAGACGGATCTAACAGTGAAGAAATGAGAAATGTGAAGGCTTCAAGAAATTAAT 1260  
DB 1313 TACTGAGACGGATCTAACAGTGAAGAAATGAGAAATGTGAAGGCTTCAAGAAATTAAT 1372  
QY 1261 ACTGTCCCAAGTGAACCTTTTAAGTGAATTCCTGAGAGATTCAGTGAAG 1320  
DB 1373 ACTGTCCCAAGTGAACCTTTTAAGTGAATTCCTGAGAGATTCAGTGAAG 1432  
QY 1321 CTACAGTTGAAGGCTATTTCTTGTGTGAAGAAATGAGATGCAAGTTCATAGTCTGTT 1380  
DB 1433 CTACAGTTGAAGGCTATTTCTTGTGTGAAGAAATGAGATGCAAGTTCATAGTCTGTT 1492  
QY 1381 AAGTCTCTAGTAAAGACATACATCACTAATAAACAAGAGATGAAGAAATTAAGTGGGA 1440  
DB 1493 AAGTCTCTAGTAAAGACATACATCACTAATAAACAAGAGATGAAGAAATTAAGTGGGA 1552  
QY 1441 TCGCTTTTGAAGTGTGTGTAGTGAACCAAGATTTGAAGAGTAACTATATGTA 1500  
DB 1553 TCGCTTTTGAAGTGTGTGTAGTGAACCAAGATTTGAAGAGTAACTATATGTA 1612  
QY 1501 GTATCCAGGGACAGTGTGTGTGTAGAGAAACAAATTCACAGATGTTCTTTAAACA 1560  
DB 1613 GTATCCAGGGACAGTGTGTGTGTAGAGAAACAAATTCACAGATGTTCTTTAAACA 1672  
QY 1561 CGAGAAATCTGTGACCTCAAAAGCCGTGTGAATTTGATTAATGAAGATGAAGTGG 1620  
DB 1673 CGAGAAATCTGTGACCTCAAAAGCCGTGTGAATTTGATTAATGAAGATGAAGTGG 1732  
QY 1621 GAAATTTAATAGTATGTTCTTAAATTCCTGTTCAGCTGTGTTTAAATTAAGTAAAG 1680  
DB 1733 GAAATTTAATAGTATGTTCTTAAATTCCTGTTCAGCTGTGTTTAAATTAAGTAAAG 1792  
QY 1681 CTATTTGAGATTAAGTGAAGTGAACCATCTGAAGAAATCTCTCTTGAAGATCTGTG 1740  
DB 1793 CTATTTGAGATTAAGTGAAGTGAACCATCTGAAGAAATCTCTCTTGAAGATCTGTG 1852  
QY 1741 ACACAGCTGACTCATAGTGTGGATGTACTGTGAACAAAGTGTAACTGTAGAT 1800  
DB 1853 ACACAGCTGACTCATAGTGTGGATGTACTGTGAACAAAGTGTAACTGTAGAT 1912  
QY 1801 GCTCTAATGATATTACAAATGGAAGATGTGCTCATGAGTGGAACTTTATTAACAGAGA 1860  
DB 1913 GCTCTAATGATATTACAAATGGAAGATGTGCTCATGAGTGGAACTTTATTAACAGAGA 1972  
QY 1861 TATTTATTAAGCATGTTTATGATTAATCTTTTTCAGTCTTTCAGAAATGTGAGCTGTG 1920  
DB 1973 TATTTATTAAGCATGTTTATGATTAATCTTTTTCAGTCTTTCAGAAATGTGAGCTGTG 2032  
QY 1921 TTGACAGATGCAAACTCAAGAGATTAATTTGATGTGTATTAAGTGAACAGAAAT 1980  
DB 2033 TTGACAGATGCAAACTCAAGAGATTAATTTGATGTGTATTAAGTGAACAGAAAT 2092  
QY 1981 GCTGAGAGTATTATGAGAGAAATGAAGACATTAATGTGAATATTATCATGCTTTCTTGG 2040  
DB 2093 GCTGAGAGTATTATGAGAGAAATGAAGACATTAATGTGAATATTATCATGCTTTCTTGG 2152  
QY 2041 GGTAGACGTCCACATGTCGAAAGCATTTTTCAGAGACTTGTGATTTGGCTAGACCAAC 2100  
DB 2153 GGTAGACGTCCACATGTCGAAAGCATTTTTCAGAGACTTGTGATTTGGCTAGACCAAC 2212  
QY 2101 ATGGTTTACAGATTTTACAGAAATTTGAAGTAACTGATCTGATCTATCTTCTGG 2160  
DB 2213 ATGGTTTACAGATTTTACAGAAATTTGAAGTAACTGATCTGATCTATCTTCTGG 2272  
QY 2161 GTGGCTACTGGTTTGTGATCTGTAGAGACTGTGGGCTTGTGAATTAACAATCACTCAAGTG 2220



```
|||||
Db 2273 GTGGCTACTGGTGTGATCTGAGGACCTGGGCTTGGACATACAACTACCTCAGTG 2332
Qy 2221 GAGCTCCAAAGCCTTCCAAACATTTTCTCATTTTGTGAATCTCCCTACTGTTATGAGA 2280
Db 2333 GAGCTCCAAAGCCTTCCAAACATTTTCTCATTTTGTGAATCTCCCTACTGTTATGAGA 2332
Qy 2281 GGTGAAGAATTTGCTTGGAAATATATATATCAATATTTGAAAGATGCCACTGAGGT 2340
Db 2293 GGTGAAGAATTTGCTTGGAAATATATATATCAATATTTGAAAGATGCCACTGAGGT 2452
Qy 2341 AAGTATATCATTTGAAAAAGTGAACAAATTTGATATCTTAATGACTTCAAGTAAATAT 2400
Db 2453 AAGTATATCATTTGAAAAAGTGAACAAATTTGATATCTTAATGACTTCAAGTAAATAT 2512
Qy 2401 GCCACAGGCCACAGACAGACCTTCTGGTCCAGTGAAGTGGGGCAACTGTTCTTTT 2460
Db 2513 GCCACAGGCCACAGACAGACCTTCTGGTCCAGTGAAGTGGGGCAACTGTTCTTTT 2572
Qy 2461 CCCATCAGGCCACACATCTGGAGAAAATTCCTATCAGTCAAGTCTTTCACCACT 2520
Db 2573 CCCATCAGGCCACACATCTGGAGAAAATTCCTATCAGTCAAGTCTTTCACCACT 2632
Qy 2521 GCTTCTGATGCTGTCAACCAGATGATTTTATGAAAAGCTGAAGATAGAAAAATCATAT 2580
Db 2633 GCTTCTGATGCTGTCAACCAGATGATTTTATGAAAAGCTGAAGATAGAAAAATCATAT 2692
Qy 2581 TCACAAATCATCTTATTAGACTGTGACGAAATGGCTACAGAGTACCCTGAAAACTTTG 2640
Db 2693 TCACAAATCATCTTATTAGACTGTGACGAAATGGCTACAGAGTACCCTGAAAACTTTG 2752
Qy 2641 AGTTCTCATATTTCTCTTAATACAGTGACTGCGAGTGAAGAGTTCAGATCACTGCAAT 2700
Db 2753 AGTTCTCATATTTCTCTTAATACAGTGACTGCGAGTGAAGAGTTCAGATCACTGCAAT 2812
Qy 2701 GGAATGTTCTTGGTCTTCCATCAATAGGCTTACCTGATTTGGAATGCTTAATGCG 2760
Db 2813 GGAATGTTCTTGGTCTTCCATCAATAGGCTTACCTGATTTGGAATGCTTAATGCG 2872
Qy 2761 TGTGTGAACAGAACATGATTAATTTGCTCAAAATATTATGATTTGGATTAATCTGACT 2820
Db 2873 TGTGTGAACAGAACATGATTAATTTGCTCAAAATATTATGATTTGGATTAATCTGACT 2932
Qy 2821 AAAAAAGAAACAACGACAGATTAATTTGAAAGAAAAAGCTCTTTCATTTATAGGCAAGT 2880
Db 2933 AAAAAAGAAACAACGACAGATTAATTTGAAAGAAAAAGCTCTTTCATTTATAGGCAAGT 2992
Qy 2881 TACCAAGAGAACTTCTCTATCAGAGGAGATGCTCTTCACTGCTTTGGGAATTAAT 2940
Db 2993 TACCAAGAGAACTTCTCTATCAGAGGAGATGCTCTTCACTGCTTTGGGAATTAAT 3052
Qy 2941 GACCTTCTGGAGACCTTGTGTGACGCTTTGTTTAAGATGTTTCTTGAAGCCGAT 3000
Db 3053 GACCTTCTGGAGACCTTGTGTGACGCTTTGTTTAAGATGTTTCTTGAAGCCGAT 3112
Qy 3001 CCTTACATAGATTAATGATCAGAAATGTTTACAGAAACATACATTTGGCTTAAAGACAT 3060
Db 3113 CCTTACATAGATTAATGATCAGAAATGTTTACAGAAACATACATTTGGCTTAAAGACAT 3172
Qy 3061 CAGAAATCAGACGATGAAATTTGGGATCCAGAAAGATGATTCATAGTACCTTCAAGT 3120
Db 3173 CAGAAATCAGACGATGAAATTTGGGATCCAGAAAGATGATTCATAGTACCTTCAAGT 3232
Qy 3121 GGCATTAATAAGTCCAGTATACCTTACAGCCTAATTTGTAATTTCTCTCCGCGGATATGA 3180
Db 3233 GGCATTAATAAGTCCAGTATACCTTACAGCCTAATTTGTAATTTCTCTCCGCGGATATGA 3292
Qy 3181 AAGTATCAGCTTAACATGATGTGCAAGAGTCTATCCATTTTGGAGTCTGAATTCAGT 3240
Db 3293 AAGTATCAGCTTAACATGATGTGCAAGAGTCTATCCATTTTGGAGTCTGAATTCAGT 3352
Qy 3241 AGAGGAATTTGACAAATTAATCTTACGCTTAATTAATGCAATGTCTCATGATGGGG 3300
|||||
Db 3353 AGAGGAATTTGACAAATTAATCTTACGCTTAATTAATGCAATGTCTCATGATGGGG 3412
Qy 3301 AGTCTTAAGCCAGAGAGAGCTTTGAATATATGCTACTGAGAGCAGAACAAAGGTGGC 3360
Db 3413 AGTCTTAAGCCAGAGAGAGCTTTGAATATATGCTACTGAGAGCAGAACAAAGGTGGC 3472
Qy 3361 ATGCAATTTGGGTGTATCAGAGTCCAAACCTTGTGACTCTGGGACCCAGCTCCCTG 3420
Db 3473 ATGCAATTTGGGTGTATCAGAGTCCAAACCTTGTGACTCTGGGACCCAGCTCCCTG 3532
Qy 3421 GATATGAATGACGCTATGACAGTCTGTCACTCTTACAAATTTGAGCTTCTGAG 3480
Db 3533 GATATGAATGACGCTATGACAGTCTGTGTCACTCTTACAAATTTGAGCTTCTGAG 3592
Qy 3481 GGAATCCCAATTAATGAGTGGCTTAAGCAGGCAAGAAATAGCTTGGGTGTTTGCACT 3540
Db 3593 GGAATCCCAATTAATGAGTGGCTTAAGCAGGCAAGAAATAGCTTGGGTGTTTGCACT 3652
Qy 3541 ACTCAGATACCACTGTGCTTTAAAGGCTGTGTGATTTGCAAGCTTAATGATATCA 3600
Db 3653 ACTCAGATACCACTGTGCTTTAAAGGCTGTGTGATTTGCAAGCTTAATGATATCA 3712
Qy 3601 GAAAGCAAAATATCCAGAGTGAACGCTGACGCGGCTTACGCTCAACAGTCTCT 3651
Db 3713 GAAAGCAAAATATCCAGAGTGAACGCTGACGCGGCTTACGCTCAACAGTCTCTTAAAGTTT 3772
Qy 3652 -----CTTGGCTGTATACAGCCA 3669
Db 3773 CTGATTGACACACAAACCGCTTACTCTTCAACAGCAGAGCTTGTGTGTACAGCCA 3832
Qy 3670 ATGGCAGTAAATATTTCCGCAAAATGTTTGAATTTGCTAATTTGCTCAATGTTTGA 3729
Db 3833 ATGGCAGTAAATATTTCCGCAAAATGTTTGAATTTGCTAATTTGCTCAATGTTTGA 3892
Qy 3730 TATTAATGAAAGCTTCTGGCTCTTCTAGAAAGCAAGAAATATCCAAATTAAGAAAGC 3789
Db 3893 TATTAATGAAAGCTTCTGGCTCTTCTAGAAAGCAAGAAATATCCAAATTAAGAAAGC 3952
Qy 3790 TTTGATTTAGATGTTGCTGTAAAGAAAGAAATTAAGATGATCTCATCATGATGATTTGAAT 3849
Db 3953 TTTGATTTAGATGTTGCTGTAAAGAAAGAAATTAAGATGATCTCATCATGATGATTTGAAT 4012
Qy 3850 GTGTGTAACAAGCTTTTCCGGCCCGGGTAGGAGTGGCAATGCTCTTATGAAATTAACCTTA 3909
Db 4013 GTGTGTAACAAGCTTTTCCGGCCCGGGTAGGAGTGGCAATGCTCTTATGAAATTAACCTTA 4072
Qy 3910 TTTAAGTGGCTTATAGTGTCTTCAAGAAACATTTCTGTGACGAGACAGTGAAGAAAGTG 3969
Db 4073 TTTAAGTGGCTTATAGTGTCTTCAAGAAACATTTCTGTGACGAGACAGTGAAGAAAGTG 4132
Qy 3970 GAATATGATCATGGAAGAACTCAACTATTTAGATTTCTGTAATGTAAGAAACCAAGTTTGT 4029
Db 4133 GAATATGATCATGGAAGAACTCAACTATTTAGATTTCTGTAATGTAAGAAACCAAGTTTGT 4192
Qy 4030 GTTAATATTTCTGTGTGAGAAAATTTAAAGTTTCAATATCCCAAGATGCTTCACTGCTCC 4089
Db 4193 GTTAATATTTCTGTGTGAGAAAATTTAAAGTTTCAATATCCCAAGATGCTTCACTGCTCC 4252
Qy 4090 ATATGATATTAATATGAGCCAAAGAGACAGGCGGTGAGAAATTAACAATCTGAAATGAAG 4149
Db 4253 ATATGATATTAATATGAGCCAAAGAGACAGGCGGTGAGAAATTAACAATCTGAAATGAAG 4312
Qy 4150 CTGTCTCTCTGTACCTTTGACATGATGTCAAGAGGCTGCGCTCTTGTGAGATGAGACT 4209
Db 4313 CTGTCTCTCTGTACCTTTGACATGATGTCAAGAGGCTGCGCTCTTGTGAGATGAGACT 4372
Qy 4210 TCAAGCTCCCATCATCACTCTTCACTATTTATTTTCTGTTTCAAGCTTCTGTACTTT 4269
Db 4373 TCAAGCTCCCATCATCACTCTTCACTATTTATTTTCTGTTTCAAGCTTCTGTACTTT 4432
Qy 4270 ATGGAATTTGGCTGTGA 4287
Db 4433 ATGGAATTTGGCTGTGA 4450
```



RESULT 8  
 AAL49815  
 AAL49815 standard; cDNA; 4335 BP.  
 AC AAL49815;  
 DT 04-DEC-2002 (first entry)  
 DE Human platelet alloantigen Gova coding sequence.  
 XX Human; platelet alloantigen; Gova; Govb; single nucleotide polymorphism;  
 XX SNP; biallelic; bleeding disorder; post-transfusion purpura;  
 XX post-transfusion platelet refractoriness; haemostatic; vaccine;  
 XX neonatal alloimmune thrombocytopenia; gene; ss.  
 OS Homo sapiens.  
 XX  
 XX Key Location/Qualifiers  
 FH 1..4335  
 FT CDS /\*tag= a  
 FT /product= "Gova"  
 FT /partial  
 FT /note= "no stop codon"  
 FT replace(2108,C)  
 FT variation /\*tag= b  
 XX  
 XX MO200270738-A2.  
 XX  
 XX 12-SEP-2002.  
 XX  
 XX 07-MAR-2002; 2002WO-CA00291.  
 XX  
 XX 07-MAR-2001; 2001US-273941P.  
 XX  
 XX (SCHU/) SCHUH A.  
 XX  
 XX Schuh A, Ouwehand W;  
 XX  
 XX MPI: 2002-713460/77.  
 XX P-PSDB; AAO19372.  
 XX  
 XX New isolated oligonucleotide binding to a region of CD109 nucleic acid  
 XX having a single nucleotide polymorphism that distinguishes a Gova  
 XX and/or Govb allele, useful for treating blood disorders e.g. alloimmune  
 XX thrombocytopenia  
 XX  
 XX Claim 6; Page 23-29; 69pp; English.  
 XX  
 XX The present invention relates to a sequence capable of binding  
 XX specifically to a CD109 nucleic acid which has a single nucleotide  
 XX polymorphism that distinguishes the Gova and Govb alleles. Detection of  
 XX the Gov genotype is useful for detecting whether the subject has or is at  
 XX risk of a blood disease, disorder or abnormal physical state, such as  
 XX bleeding, or increased risk of bleeding, due to alloimmune destruction of  
 XX blood platelets, e.g., post-transfusion purpura, post-transfusion  
 XX platelet refractoriness or neonatal alloimmune thrombocytopenia. The  
 XX nucleic acid and polypeptide are useful for Gov genotyping or phenotyping  
 XX individuals. The present sequence is the Gova coding sequence.  
 XX  
 XX Sequence 4335 BP; 1296 A; 837 C; 922 G; 1280 T; 0 other;  
 XX  
 XX Query Match 98.5%; Score 4221.4; DB 24; Length 4335;  
 XX Best Local Similarity 98.8%; Pred. No. 0;  
 XX Matches 4283; Conservative 0; Mismatches 1; Indels 51; Gaps 1;

QY 1 ATGAGAGGCGCCAGCTCTGAGACCGCGCCACCTCTGAGTGAGACCGCGGCTG 60  
 DB 1 ATGAGAGGCGCCAGCTCTGAGACCGCGCCACCTCTGAGTGAGACCGCGGCTG 60  
 QY 61 GCGTGAGCTCCGCGGCTCGGTTCTGTGAGACGCCCAAGGATCATCAGGCCCGAGGA 120  
 DB 1141 ACTTTGAAGAAAGAAATATGTAGTCAATACAGTACAGAGAAACTATATCTGAG 1200  
 DB 1141 ACTTTGAAGAAAGAAATATGTAGTCAATACAGTACAGAGAAACTATATCTGAG 1200

DB 61 GCGTGAGCTCCGCGGCTCGGTTCTGTGAGACGCCCAAGGATCATCAGGCCCGAGGA 120  
 QY 121 AATGACTATTTGGGGTGGAGCTTCTGGAAACACTGCCCTTCAAGGTAGTGAAGCG 180  
 DB 121 AATGACTATTTGGGGTGGAGCTTCTGGAAACACTGCCCTTCAAGGTAGTGAAGCG 180  
 QY 181 GAGCTGCTCAAGACAGCATCAAACTCACTGCTCTGTCTGTGAGAGCAGAAAGACTCTT 240  
 DB 181 GAGCTGCTCAAGACAGCATCAAACTCACTGCTCTGTCTGTGAGAGCAGAAAGACTCTT 240  
 QY 241 GAAAAAGGCTCTTTTAAGACACTTACTCTTCATCTACTACTCTGTAACAGTSCAATGAG 300  
 DB 241 GAAAAAGGCTCTTTTAAGACACTTACTCTTCATCTACTACTCTGTAACAGTSCAATGAG 300  
 QY 301 ATTATGAGTACGCTGTAACCGGACGTAACCGAGTGAATTTTATCTCTATATGTAAC 360  
 DB 301 ATTATGAGTACGCTGTAACCGGACGTAACCGAGTGAATTTTATCTCTATATGTAAC 360  
 QY 361 CGCTTATCATTTGAGACCAAGAAATATCTGCTTCAATCAACAGACAGGCTTATAC 420  
 DB 361 CGCTTATCATTTGAGACCAAGAAATATCTGCTTCAATCAACAGACAGGCTTATAC 420  
 QY 421 AAGCCAAAGCAAGAGTGAATTTGCGCATTTGTAACCTCTCTCAAGTTTAAAGCTTAC 480  
 DB 421 AAGCCAAAGCAAGAGTGAATTTGCGCATTTGTAACCTCTCTCAAGTTTAAAGCTTAC 480  
 QY 481 AAAACCTCTTTAAACATTTCTAATTAAGACCCCAAAATTTGATCCAAAGGTTG 540  
 DB 481 AAAACCTCTTTAAACATTTCTAATTAAGACCCCAAAATTTGATCCAAAGGTTG 540  
 QY 541 TCACAACAAAGTATCTTGAAGTCAATTTCAAACTTTTCAAGTATCTTCCATCAATA 600  
 DB 541 TCACAACAAAGTATCTTGAAGTCAATTTCAAACTTTTCAAGTATCTTCCATCAATA 600  
 QY 601 CTGTGAGCTGCTATTTCAAGTCAAGTGAATGACCAAGATATTAATCATTTAG 660  
 DB 601 CTGTGAGCTGCTATTTCAAGTCAAGTGAATGACCAAGATATTAATCATTTAG 660  
 QY 661 GTTTCAGAAATGATTAACCAAAATTTGAAGGACTTTGACAGACCAATTAATGTTCT 720  
 DB 661 GTTTCAGAAATGATTAACCAAAATTTGAAGGACTTTGACAGACCAATTAATGTTCT 720  
 QY 721 ATGAATTTAAGCATTTAATGATGATCACTCAAGCAAGATATCATTTAGGAGCACTG 780  
 DB 721 ATGAATTTAAGCATTTAATGATGATCACTCAAGCAAGATATCATTTAGGAGCACTG 780  
 QY 781 AAAGGAGCGTAAAGCTTACATTTTACCTTTTACCTTTTGGGGAAGAAAGAAATAT 840  
 DB 781 AAAGGAGCGTAAAGCTTACATTTTACCTTTTACCTTTTGGGGAAGAAAGAAATAT 840  
 QY 841 ACAAACAACTTAAATGATTAATGATGATCACTCAAGCAAGATATCATTTAGGAGCACTG 900  
 DB 841 ACAAACAACTTAAATGATTAATGATGATCACTCAAGCAAGATATCATTTAGGAGCACTG 900  
 QY 901 AATGATGATGATTTCTTCAATGAGACTTTCTGAAATCTGATCTCTTCCCTGAGCA 960  
 DB 901 AATGATGATGATTTCTTCAATGAGACTTTCTGAAATCTGATCTCTTCCCTGAGCA 960  
 QY 961 GTAGAAATTTTAAACCAAGTACAGAAATCACTTACAGGATTTTCAAGAAATGTAAGCACT 1020  
 DB 961 GTAGAAATTTTAAACCAAGTACAGAAATCACTTACAGGATTTTCAAGAAATGTAAGCACT 1020  
 QY 1021 AATGATTTCTTCAAGCAAGATGATTAATGATGATTTTGAATTAATCACTGCTTGG 1080  
 DB 1021 AATGATTTCTTCAAGCAAGATGATTAATGATGATTTTGAATTAATCACTGCTTGG 1080  
 QY 1081 AAGCATCTCTCAACTTCAAGCAAGCTGTAAGGTAATCTGCTGATGAGCAACCACTG 1140  
 DB 1081 AAGCATCTCTCAACTTCAAGCAAGCTGTAAGGTAATCTGCTGATGAGCAACCACTG 1140  
 QY 1141 ACTTTGAAGAAAGAAATATGTAGTCAATACAGTACAGAGAAACTATATCTGAG 1200  
 DB 1141 ACTTTGAAGAAAGAAATATGTAGTCAATACAGTACAGAGAAACTATATCTGAG 1200

QY	1201	TACTGAGCCGGATCTTAACAAGTGGAAATCAGAAATGAGAGCTGTCAGAAATATAATAT	1260
Db	1201	TACTGAGACGGATCTTAACAAGTGGAAATCAGAAATGAGAGCTGTCAGAAATATAATAT	1260
QY	1261	ACTGCCCCCAAGGGAACCTTTTAAGTTGAAATCCCAATCCCTGGAGAGATCCAGTGAG	1320
Db	1261	ACTGCCCCCAAGGGAACCTTTTAAGTTGAAATCCCAATCCCTGGAGAGATCCAGTGAG	1320
QY	1321	CTACAGTTGAAGCCATTTCTTCCTGGTAGTAAAGTAGCATGGCAGTTTCATAGTCGTGTT	1380
Db	1321	CTACAGTTGAAGCCATTTCTTCCTGGTAGTAAAGTAGCATGGCAGTTTCATAGTCGTGTT	1380
QY	1361	AAGTCTCTAGTAAGACATACATCCAACTAAACCAAGAGATGAATAATATAAGGTGGGA	1440
Db	1361	AAGTCTCTAGTAAGACATACATCCAACTAAACCAAGAGATGAATAATATAAGGTGGGA	1440
QY	1441	TCGCCCTTTGAGTTGGTGGTTAAGTGGCAACAACGATGGAAGAGTTAAAGTATATGGTA	1500
Db	1441	TCGCCCTTTGAGTTGGTGGTTAAGTGGCAACAACGATGGAAGAGTTAAAGTATATGGTA	1500
QY	1501	GTATCCAGGGGACAGTTGGTGGCTGTAGGAAAAACAAATTAACAATGTCTCTTTAACA	1560
Db	1501	GTATCCAGGGGACAGTTGGTGGCTGTAGGAAAAACAAATTAACAATGTCTCTTTAACA	1560
QY	1561	CCAGAAAAATCTTGAGACTCCAAAAGCCTGTGTAAATGTATATATTGAAGATGATGGG	1620
Db	1561	CCAGAAAAATCTTGAGACTCCAAAAGCCTGTGTAAATGTATATATTGAAGATGATGGG	1620
QY	1621	GAAATTATTAAGTGAATGTTCTTAAAAATTCCTGTTCAGCTTGTTTTTAAAAATTAAGTAAAG	1680
Db	1621	GAAATTATTAAGTGAATGTTCTTAAAAATTCCTGTTCAGCTTGTTTTTAAAAATTAAGTAAAG	1680
QY	1661	CTAATATGAGATGAAGTAAAGCTGAACCATCTGGAAGAAAGTCCTCTTAAGGATCTGCTG	1740
Db	1661	CTAATATGAGATGAAGTAAAGCTGAACCATCTGGAAGAAAGTCCTCTTAAGGATCTGCTG	1740
QY	1741	ACACAGCCTGACTCCATAGTTGGAGTTGGATGCTGTGCAAAAAGTGTGAATCTGATGAAT	1800
Db	1741	ACACAGCCTGACTCCATAGTTGGAGTTGGATGCTGTGCAAAAAGTGTGAATCTGATGAAT	1800
QY	1801	GCTCTTAATGATATTACAATGAGAAAATGTGTCATGAGTTGGAACCTTTATPAACAACAGA	1860
Db	1801	GCTCTTAATGATATTACAATGAGAAAATGTGTCATGAGTTGGAACCTTTATPAACAACAGA	1860
QY	1861	TATATATTAAGCAATGCTTCATGAATCTTTTCACAGCTTTCACGAAATGCGACTCGGTA	1920
Db	1861	TATATATTAAGCAATGCTTCATGAATCTTTTCACAGCTTTCACGAAATGCGACTCGGTA	1920
QY	1921	TTGACAGATGCAAACTCCACAGAGGATATATATGATGGTGTATATGACAAATGACGAATAT	1980
Db	1921	TTGACAGATGCAAACTCCACAGAGGATATATATGATGGTGTATATGACAAATGACGAATAT	1980
QY	1981	GCTGAGAGGTTTATGAGAGAAAATGAAAGACATATGTGAATATTCATGACTTTCTTTG	2040
Db	1981	GCTGAGAGGTTTATGAGAGAAAATGAAAGACATATGTGAATATTCATGACTTTCTTTG	2040
QY	2041	GGTAGCAATCCACATGTCCGAAAGCATTTTCACAGACTTGGACTTGAACCAAC	2100
Db	2041	GGTAGCAATCCACATGTCCGAAAGCATTTTCACAGACTTGGACTTGAACCAAC	2100
QY	2101	ATGGGTTACAGAGATTACCAAGAATTTGAAGTAACTGAACTGATCTGATCACTTCCTGG	2160
Db	2101	ATGGGTTACAGAGATTACCAAGAATTTGAAGTAACTGAACTGATCTGATCACTTCCTGG	2160
QY	2161	GTGGCTACTGGTTTTGTGATCTGTGAGGACCTGGGCTCTTGACATPAACAATCACTCAAGTG	2220
Db	2161	GTGGCTACTGGTTTTGTGATCTGTGAGGACCTGGGCTCTTGACATPAACAATCACTCAAGTG	2220
QY	2221	GAGCTCCAGGCTTCAACAATTTTCAATTTTTTTGAATCTTCCCTACTCTGTTATCAGA	2280
Db	2221	GAGCTCCAGGCTTCAACAATTTTCAATTTTTTTGAATCTTCCCTACTCTGTTATCAGA	2280

QY	2281	GGTGAAGAAATTTGCTTGGAAATACATATATCTAATTAATTTGAAAGATGCCACGTAGAGTT	2340
Db	2281	GGTGAAGAAATTTGCTTGGAAATACATATATCTAATTAATTTGAAAGATGCCACGTAGAGTT	2340
QY	2341	AAGGTAATCATTTGAGAAAGTGAACAAATTTGATATTTCTAATGACTTCAAGTGAATAAT	2400
Db	2341	AAGGTAATCATTTGAGAAAGTGAACAAATTTGATATTTCTAATGACTTCAATGAAATTAAT	2400
QY	2401	GCCAAGGCCACACAGACAGACCCCTTCTGATTCACAGTAGAGATGGGGCAACTGTTCTTTT	2460
Db	2401	GCCAAGGCCACACAGACAGACCCCTTCTGATTCACAGTAGAGATGGGGCAACTGTTCTTTT	2460
QY	2461	CCCATCAGGCCAACACATCTGGGAGAAATTCCTATCAAGTCAACAGCTCTTTCACCACT	2520
Db	2461	CCCATCAGGCCAACACATCTGGGAGAAATTCCTATCAAGTCAACAGCTCTTTCACCACT	2520
QY	2521	GCTCTGATGCTGCACCCAGATATATTTAGTAAAGCTGAAGGAATAGAAATATCATAT	2580
Db	2521	GCTCTGATGCTGCACCCAGATATATTTAGTAAAGCTGAAGGAATAGAAATATCATAT	2580
QY	2581	TCACAATCCATCTTATTAGACTTGACTGACATAGGCTACAGAGTACCTGAAAACCTTGG	2640
Db	2581	TCACAATCCATCTTATTAGACTTGACTGACATAGGCTACAGAGTACCTGAAAACCTTGG	2640
QY	2641	AGTTTCTCATTTCTCTTAATACGTGACTGGCAGTGAAGAAGTTCAGATCACTGCATTT	2700
Db	2641	AGTTTCTCATTTCTCTTAATACGTGACTGGCAGTGAAGAAGTTCAGATCACTGCATTT	2700
QY	2701	GGAGATGCTTGGTCCCTTCATCAATGGCTTAGCCCATTTGANTGGATGCCATTAGGC	2760
Db	2701	GGAGATGCTTGGTCCCTTCATCAATGGCTTAGCCCATTTGANTGGATGCCATTAGGC	2760
QY	2761	TGTGGTGAACGAACATGATTAATTTTGTCTCCAAATATTTTACATTTTGGATTTATCTGACT	2820
Db	2761	TGTGGTGAACGAACATGATTAATTTTGTCTCCAAATATTTTACATTTTGGATTTATCTGACT	2820
QY	2821	AAAAAGAAACAACCTGACAGATTAATTTGAAAAGAAAGCTCTTCATTTATGAGGCAGAGT	2880
Db	2821	AAAAAGAAACAACCTGACAGATTAATTTGAAAAGAAAGCTCTTCATTTATGAGGCAGAGT	2880
QY	2881	TACCAAGAGAACTTCTCTATCAGAGGGAAGAGGCTCTTCAGTGGCTTTTGGAAATAT	2940
Db	2881	TACCAAGAGAACTTCTCTATCAGAGGGAAGAGGCTCTTCAGTGGCTTTTGGAAATAT	2940
QY	2941	GACCCCTCGGAGACACTTGGTGTGACGCTTTTGTTTAAGATGTTTCTCTTGAAGCCGAT	3000
Db	2941	GACCCCTCGGAGACACTTGGTGTGACGCTTTTGTTTAAGATGTTTCTCTTGAAGCCGAT	3000
QY	3001	CCTTACATAGATATTGATCAGAAATGTGTTACACAGAAATATACACTTGGCTTAAAGACAT	3060
Db	3001	CCTTACATAGATATTGATCAGAAATGTGTTACACAGAAATATACACTTGGCTTAAAGACAT	3060
QY	3061	CAGAAATCCAAACGGTGAATTTTGGGATCCAGAGAGAGTGAATTCATAGTAGCTTCAAGGT	3120
Db	3061	CAGAAATCCAAACGGTGAATTTTGGGATCCAGAGAGAGTGAATTCATAGTAGCTTCAAGGT	3120
QY	3121	GGCAATAAAGTCCAGTACACTTACAGCCCTATATTGTAACCTTCTCCCTGGGATATAGA	3180
Db	3121	GGCAATAAAGTCCAGTACACTTACAGCCCTATATTGTAACCTTCTCTCCCTGGGATATAGA	3180
QY	3181	AAGTATCAGCCTTAACATTTGATGTGCAAGAGTCTATCCACTTTTGGAGTCTGAATTCAGT	3240
Db	3181	AAGTATCAGCCTTAACATTTGATGTGCAAGAGTCTATCCACTTTTGGAGTCTGAATTCAGT	3240
QY	3241	AGAGGAATTTCAACAATTAATCTTAGCCCTTAATACTTATGATTTGTATCATCAGTGGGG	3300
Db	3241	AGAGGAATTTCAACAATTAATCTTAGCCCTTAATACTTATGATTTGTATCATCAGTGGGG	3300
QY	3301	AGTCTTAAGGGAAGAGGCTTGAATATGCTGACTTGGAGAGCAGAAACAAGAGTGGC	3360
Db	3301	AGTCTTAAGGGAAGAGGCTTGAATATGCTGACTTGGAGAGCAGAAACAAGAGTGGC	3360
QY	3361	ATGCAATTTGGGATGATCAGAGTCCAAACTTCTCTACTCTGGCAGCCACGCTCTCCTG	3420



181 GAGTGTCAAGACGATCAAACTCACTGTCTGTCTGAGACGAAAGAGTCTT 240  
241 GAAAAAGGCTCTTTTAAGACATTACTTCTCATCACTACCTCTGAACAGTGCAGATGAG 300  
241 GAAAAAGGCTCTTTTAAGACATTACTTCTCATCACTACCTCTGAACAGTGCAGATGAG 300  
301 ATTTATGAGCTACGTGTAAACCGAGATACCCAGAGTAGATTTTATCTCTAATAGTACC 360  
301 ATTTATGAGCTACGTGTAAACCGAGATACCCAGAGTAGATTTTATCTCTAATAGTACC 360  
361 CGCTTATATTTGAGACCAAGAGATATCTGTCTTCAATTCAAACAGACAGGCTTATAC 420  
361 CGCTTATATTTGAGACCAAGAGATATCTGTCTTCAATTCAAACAGACAGGCTTATAC 420  
421 AAGCCAAAGCAAGAGTAGATTTGAGATTTGATGATTCACAGTGTG 540  
421 AAGCCAAAGCAAGAGTAGATTTGAGATTTGATGATTCACAGTGTG 540  
481 AAAACCTCTTTAAACATTTCTCATTAAGGACCCCAATCAATTTGATTCACAGTGTG 540  
481 AAAACCTCTTTAAACATTTCTCATTAAGGACCCCAATCAATTTGATTCACAGTGTG 540  
541 TCACAAACAAAGTATCTTGAGTCAATTCGAAAATTTTCAAGTATCTTCCATCCATA 600  
541 TCACAAACAAAGTATCTTGAGTCAATTCGAAAATTTTCAAGTATCTTCCATCCATA 600  
601 CTGTGTGCTGTGCTATTCAGAGTCAAGTGAATGACGACATTTATCATTCATTCAG 660  
601 CTGTGTGCTGTGCTATTCAGAGTCAAGTGAATGACGACATTTATCATTCATTCAG 660  
661 GTTTCAGAAATATGATTAACCAAAATTTGAGTGAATTTGACAGACCATTAATTTGTTCT 720  
661 GTTTCAGAAATATGATTAACCAAAATTTGAGTGAATTTGACAGACCATTAATTTGTTCT 720  
721 ATGAAATTTAAGCATTTAATGATGATCACTACGCGAAAGTATACATATGGAAGCCAGTG 780  
721 ATGAAATTTAAGCATTTAATGATGATCACTACGCGAAAGTATACATATGGAAGCCAGTG 780  
781 AAGAGAGAGTGAAGCTTACATTTTACCTTTATCTTTTGGGAAAGAAATTAATTT 840  
781 AAGAGAGAGTGAAGCTTACATTTTACCTTTATCTTTTGGGAAAGAAATTAATTT 840  
841 ACAAACCATTTAAGATTAATGATGATCTGCAAACTTCTTTTAAATGATGAAGATGAA 900  
841 ACAAACCATTTAAGATTAATGATGATCTGCAAACTTCTTTTAAATGATGAAGATGAA 900  
901 AATGTAATGATTTCTTCAATGAGACTTTCTGAATACCTGATCTATCTTCCCTGAGACA 960  
901 AATGTAATGATTTCTTCAATGAGACTTTCTGAATACCTGATCTATCTTCCCTGAGACA 960  
961 GTAGAAATTTTAAACCAAGTGAAGATCACTGATCAAGTATTTTAAAGAAATGTAAGACT 1020  
961 GTAGAAATTTTAAACCAAGTGAAGATCACTGATCAAGTATTTTAAAGAAATGTAAGACT 1020  
1021 AATGTTCTTCAAGCAATGATTAATCATGATGATTTTGAATTAATCACTAGTCTTG 1080  
1021 AATGTTCTTCAAGCAATGATTAATCATGATGATTTTGAATTAATCACTAGTCTTG 1080  
1081 AAGCATCTCTCAACTTCACAGCACTGTGAAGGTAATCTGTCTGATGAGCAACCACTG 1140  
1081 AAGCATCTCTCAACTTCACAGCACTGTGAAGGTAATCTGTCTGATGAGCAACCACTG 1140  
1141 ACTTTGAAGAAAGAAATTAATGATGATCACTAATCAGTGAACAGAGAAATTAATCTGAG 1200  
1141 ACTTTGAAGAAAGAAATTAATGATGATCACTAATCAGTGAACAGAGAAATTAATCTGAG 1200  
1201 TACTGAGCGGATCTAATCAGTGAAGATGATGAAGATGATGATGATGATGATGATGAT 1260  
1201 TACTGAGCGGATCTAATCAGTGAAGATGATGAAGATGATGATGATGATGATGATGAT 1260  
1261 ACTGTCCCAAGTGAAGCTTTTAAATGATGATGATGATGATGATGATGATGATGATGAT 1320  
1261 ACTGTCCCAAGTGAAGCTTTTAAATGATGATGATGATGATGATGATGATGATGATGAT 1320

1261 ACTGTCCCAAGTGAAGCTTTTAAATGATGATGATGATGATGATGATGATGATGATGAT 1320  
1321 CTACAGTTGAAGGCTATTTCTTTGATGATGATGATGATGATGATGATGATGATGATGAT 1380  
1321 CTACAGTTGAAGGCTATTTCTTTGATGATGATGATGATGATGATGATGATGATGATGAT 1380  
1381 AAGTCTCTTGAAGATCATTCACATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1440  
1381 AAGTCTCTTGAAGATCATTCACATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1440  
1441 TCGCTTTTGAAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1500  
1441 TCGCTTTTGAAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1500  
1501 GTATCCAGGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1560  
1501 GTATCCAGGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1560  
1561 CCAAGAAATTTCTGAGACTCCAAAGGCTGTGTATGATGATGATGATGATGATGATGATGAT 1620  
1561 CCAAGAAATTTCTGAGACTCCAAAGGCTGTGTATGATGATGATGATGATGATGATGATGAT 1620  
1621 GAAATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1680  
1621 GAAATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1680  
1681 CTATATTTGAGATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1740  
1681 CTATATTTGAGATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1740  
1741 ACAAGCTGATCTCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1800  
1741 ACAAGCTGATCTCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1800  
1801 GCTCTTAATGATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1860  
1801 GCTCTTAATGATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1860  
1861 TATTAATTTAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1920  
1861 TATTAATTTAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1920  
1921 TTGACAGATGCAAACTTCACAGAGATTAATGATGATGATGATGATGATGATGATGATGAT 1980  
1921 TTGACAGATGCAAACTTCACAGAGATTAATGATGATGATGATGATGATGATGATGATGAT 1980  
1981 GCTGAGAGTTTGAAGAGAAATGAGAGATTAATGATGATGATGATGATGATGATGATGATGAT 2040  
1981 GCTGAGAGTTTGAAGAGAAATGAGAGATTAATGATGATGATGATGATGATGATGATGATGAT 2040  
2041 GTAGAGATGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2100  
2041 GTAGAGATGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2100  
2101 ATGGGTTTCAAGATTTTACCAAGATTTGAAGTATGATGATGATGATGATGATGATGATGAT 2160  
2101 ATGGGTTTCAAGATTTTACCAAGATTTGAAGTATGATGATGATGATGATGATGATGATGAT 2160  
2161 GTGGCTTACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2220  
2161 GTGGCTTACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2220  
2221 GAGCTCCAGGCTTCCACATTTTCAATTTTGAATTTTGAATTTTGAATTTTGAATTTTGAAT 2280  
2221 GAGCTCCAGGCTTCCACATTTTCAATTTTGAATTTTGAATTTTGAATTTTGAATTTTGAAT 2280  
2281 GGTGAAGATTTGCTTTGGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2340  
2281 GGTGAAGATTTGCTTTGGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2340  
2341 GGTGAAGATTTGCTTTGGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2400  
2341 GGTGAAGATTTGCTTTGGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2400

QY 2401 GCCACAGGCCACGACGACCTTCTGTTCCAGTAGAGATGGGCAACTGTTCTTTT 2460  
Db 2401 GCCACAGGCCACGACGACCTTCTGTTCCAGTAGAGATGGGCAACTGTTCTTTT 2460  
QY 2461 CCCATAGGCGCAACATCTGGGAGAAATTCCTATCAGATCAGCTCTTCAACCCT 2520  
Db 2461 CCCATAGGCGCAACATCTGGGAGAAATTCCTATCAGATCAGCTCTTCAACCCT 2520  
QY 2521 GCTTCTGATGCTGACCCGATGATTTTATGTAAGGCTGAAGAAATAGAAATCATAT 2580  
Db 2521 GCTTCTGATGCTGACCCGATGATTTTATGTAAGGCTGAAGAAATAGAAATCATAT 2580  
QY 2581 TCACATTCATCTTATTAGACTGACGAAATAGGCTACAGATACCTGAAACCTTGG 2640  
Db 2581 TCACATTCATCTTATTAGACTGACGAAATAGGCTACAGATACCTGAAACCTTGG 2640  
QY 2641 AGTTTCTCATTTCTCTCTATACAGTACGCGCAGTGAAGAAAGATTGATCACTCAAT 2700  
Db 2641 AGTTTCTCATTTCTCTCTATACAGTACGCGCAGTGAAGAAAGATTGATCACTCAAT 2700  
QY 2701 GGAGATGTTCTTGCTCTTCCATCAATGGCTTACGCTCATGTAATGGATGCTTATGGC 2760  
Db 2701 GGAGATGTTCTTGCTCTTCCATCAATGGCTTACGCTCATGTAATGGATGCTTATGGC 2760  
QY 2761 TGTGTGAACAGACATGATAATTTTGTCTCAAAATATTTTCAATTTGATTAATGACT 2820  
Db 2761 TGTGTGAACAGACATGATAATTTTGTCTCAAAATATTTTCAATTTGATTAATGACT 2820  
QY 2821 AAAAAGAACACTGACAGATAATTTGAAAGAAAGAAAGCTTTTCAATTTATGAGCAAGT 2880  
Db 2821 AAAAAGAACACTGACAGATAATTTGAAAGAAAGAAAGCTTTTCAATTTATGAGCAAGT 2880  
QY 2881 TACAGAGAACTTCTATCAGAGGAAAGTGTCTTCTGATGCTTTTGGAAATTAAT 2940  
Db 2881 TACAGAGAACTTCTATCAGAGGAAAGTGTGTCTTCTGATGCTTTTGGAAATTAAT 2940  
QY 2941 GACCTTCTGGAGCACTTGTGTGAGCTTTTGTAAAGATGTTTCTTGAAGCCGAT 3000  
Db 2941 GACCTTCTGGAGCACTTGTGTGAGCTTTTGTAAAGATGTTTCTTGAAGCCGAT 3000  
QY 3001 CCTTACATGATATTGATCAGAAATGTTTACAGAAACATACCTTGGCTTAAAGACAT 3060  
Db 3001 CCTTACATGATATTGATCAGAAATGTTTACAGAAACATACCTTGGCTTAAAGACAT 3060  
QY 3061 CAGAAATCCAGCGTGAATTTTGGGATCCAGGAAAGTGAATTAATGAGCTTCAAGT 3120  
Db 3061 CAGAAATCCAGCGTGAATTTTGGGATCCAGGAAAGTGAATTAATGAGCTTCAAGT 3120  
QY 3121 GGCATTAAGATCCAGTAACCTTACAGGCTATATGTAATCTCTCCGCGGATATAGA 3180  
Db 3121 GGCATTAAGATCCAGTAACCTTACAGGCTATATGTAATCTCTCCGCGGATATAGA 3180  
QY 3181 AAGTATCAGCTTAACATGTAATGCAAGATCTATTCATTTTGGAGTCTGAATCAGT 3240  
Db 3181 AAGTATCAGCTTAACATGTAATGCAAGATCTATTCATTTTGGAGTCTGAATCAGT 3240  
QY 3241 AAGGGAATTTCAACAATTAATCTAGGCTTATTAATCTTATGATGATGAGGAG 3300  
Db 3241 AAGGGAATTTCAACAATTAATCTAGGCTTATTAATCTTATGATGATGAGGAG 3300  
QY 3301 AGTCTTAAGCGAAGGAGCTTGAATATCTGACCTTGAAGACAAACAGAAAGGTGGC 3360  
Db 3301 AGTCTTAAGCGAAGGAGCTTGAATATGCTGACCTTGAAGACAAACAGAAAGGTGGC 3360  
QY 3361 ATGCAATTTCTGGTGTATCAGAGTCCAACTTTGACTCTCGGACGCAAGCTCCCTG 3420  
Db 3361 ATGCAATTTCTGGTGTATCAGAGTCCAACTTTGACTCTCGGACGCAAGCTCCCTG 3420  
QY 3421 GATATTGAAGTTGACGCTTATGCACTGCTTCACTTTTCAATTTTCAAGCTTCTGAG 3480  
Db 3421 GATATTGAAGTTGACGCTTATGCACTGCTTCACTTTTCAATTTTCAAGCTTCTGAG 3480

QY 3481 GGAATCCCAATTAAGAGTGGCTTAAGCAGGCAAGAAATAGCTGGTGGTTTGCATCT 3540  
Db 3481 GGAATCCCAATTAAGAGTGGCTTAAGCAGGCAAGAAATAGCTGGTGGTTTGCATCT 3540  
QY 3541 ACTCAGATACCACTGTGGCTTTAAAGGCTGTGTGTAATTTGAGCCCTAATGAATCA 3600  
Db 3541 ACTCAGATACCACTGTGGCTTTAAAGGCTGTGTGTAATTTGAGCCCTAATGAATCA 3600  
QY 3601 GAAAGGACAAATATCAAGTGAACGCTGACGGGGCTAGCTGACCAAGTCTT----- 3651  
Db 3601 GAAAGGACAAATATCAAGTGAACGCTGACGGGGCTAGCTGACCAAGTCTTAAAGTTT 3660  
QY 3652 -----CTTCTGTGTACAGCCA 3669  
Db 3652 CTGATTTGACACACACACACCGCTTACTCTTACAGACAGAGCTTGTGTGTAACGCCA 3720  
QY 3670 ATGGAGTTAATATTTCCGCAATGCTTTTGGATTTGCTAATTTGCACTCAATGTTGA 3729  
Db 3670 ATGGAGTTAATATTTCCGCAATGCTTTTGGATTTGCTAATTTGCACTCAATGTTGA 3729  
QY 3721 ATGGAGTTAATATTTCCGCAATGCTTTTGGATTTGCTAATTTGCACTCAATGTTGA 3780  
Db 3721 ATGGAGTTAATATTTCCGCAATGCTTTTGGATTTGCTAATTTGCACTCAATGTTGA 3780  
QY 3730 TATAATGGAAGGCTTCTGGGCTTCTGAAAGAGAAATCTATCCAAATCAAGAGCC 3789  
Db 3730 TATAATGGAAGGCTTCTGGGCTTCTGAAAGAGAAATCTATCCAAATCAAGAGCC 3789  
QY 3781 TATAATGGAAGGCTTCTGGGCTTCTGAAAGAGAAATCTATCCAAATCAAGAGCC 3840  
Db 3781 TATAATGGAAGGCTTCTGGGCTTCTGAAAGAGAAATCTATCCAAATCAAGAGCC 3840  
QY 3790 TTTGATTTAGATGTTGCTGTAAGAAATTAAGATGATCTCAATCATGTGATTTGAAT 3849  
Db 3790 TTTGATTTAGATGTTGCTGTAAGAAATTAAGATGATCTCAATCATGTGATTTGAAT 3849  
QY 3841 TTTGATTTAGATGTTGCTGTAAGAAATTAAGATGATCTCAATCATGTGATTTGAAT 3900  
Db 3841 TTTGATTTAGATGTTGCTGTAAGAAATTAAGATGATCTCAATCATGTGATTTGAAT 3900  
QY 3850 GTGTGTACAGCTTTTGGGCGCGGCTGAGAGTGCATGCTCTTATGGAATTAACCTA 3909  
Db 3850 GTGTGTACAGCTTTTGGGCGCGGCTGAGAGTGCATGCTCTTATGGAATTAACCTA 3909  
QY 3901 GTGTGTACAGCTTTTGGGCGCGGCTGAGAGTGCATGCTCTTATGGAATTAACCTA 3960  
Db 3901 GTGTGTACAGCTTTTGGGCGCGGCTGAGAGTGCATGCTCTTATGGAATTAACCTA 3960  
QY 3910 TTAAGTGTATTAAGTGTGCTTCAAGACAAATTTCTGAGGAGACAGTGAAGAAAGT 3969  
Db 3910 TTAAGTGTATTAAGTGTGCTTCAAGACAAATTTCTGAGGAGACAGTGAAGAAAGT 3969  
QY 3961 TTAAGTGTATTAAGTGTGCTTCAAGACAAATTTCTGAGGAGACAGTGAAGAAAGT 4020  
Db 3961 TTAAGTGTATTAAGTGTGCTTCAAGACAAATTTCTGAGGAGACAGTGAAGAAAGT 4020  
QY 3970 GAATATGATCATGAGAACTCAACCTTATTTAGATCTGTAATGAATGAACCAAGTTTGT 4029  
Db 3970 GAATATGATCATGAGAACTCAACCTTATTTAGATCTGTAATGAATGAACCAAGTTTGT 4029  
QY 4021 GAATATGATCATGAGAACTCAACCTTATTTAGATCTGTAATGAATGAACCAAGTTTGT 4080  
Db 4021 GAATATGATCATGAGAACTCAACCTTATTTAGATCTGTAATGAATGAACCAAGTTTGT 4080  
QY 4030 GTTAATATTTCTGCTGTGAGAAACTTTAAAGTTTCAAAATCCAGATGCTTCACTGTC 4089  
Db 4030 GTTAATATTTCTGCTGTGAGAAACTTTAAAGTTTCAAAATCCAGATGCTTCACTGTC 4089  
QY 4081 GTTAATATTTCTGCTGTGAGAAACTTTAAAGTTTCAAAATCCAGATGCTTCACTGTC 4140  
Db 4081 GTTAATATTTCTGCTGTGAGAAACTTTAAAGTTTCAAAATCCAGATGCTTCACTGTC 4140  
QY 4090 ATAGTGAATTAATAGACCAAGAGACAGCGGCTGAGAAATTAATCTGTAATGAAT 4149  
Db 4090 ATAGTGAATTAATAGACCAAGAGACAGCGGCTGAGAAATTAATCTGTAATGAAT 4149  
QY 4141 ATAGTGAATTAATAGACCAAGAGACAGCGGCTGAGAAATTAATCTGTAATGAAT 4200  
Db 4141 ATAGTGAATTAATAGACCAAGAGACAGCGGCTGAGAAATTAATCTGTAATGAAT 4200  
QY 4150 CTGTCTCTGTGACCTTGTGACAGTATGCCAGGCTGCCGCTTGTGAGGATGAGCT 4209  
Db 4150 CTGTCTCTGTGACCTTGTGACAGTATGCCAGGCTGCCGCTTGTGAGGATGAGCT 4209  
QY 4201 CTGTCTCTGTGACCTTGTGACAGTATGCCAGGCTGCCGCTTGTGAGGATGAGCT 4260  
Db 4201 CTGTCTCTGTGACCTTGTGACAGTATGCCAGGCTGCCGCTTGTGAGGATGAGCT 4260  
QY 4210 TCAGGCTCCCATCATCACTCTTCACTCAATTTTATTTCTGTTCAAGCTTCTGACTTT 4269  
Db 4210 TCAGGCTCCCATCATCACTCTTCACTCAATTTTATTTCTGTTCAAGCTTCTGACTTT 4269  
QY 4261 TCAGGCTCCCATCATCACTCTTCACTCAATTTTATTTCTGTTCAAGCTTCTGACTTT 4320  
Db 4261 TCAGGCTCCCATCATCACTCTTCACTCAATTTTATTTCTGTTCAAGCTTCTGACTTT 4320  
QY 4270 ATGGAATTTGGCTG 4284  
Db 4270 ATGGAATTTGGCTG 4284  
QY 4321 ATGGAATTTGGCTG 4335  
Db 4321 ATGGAATTTGGCTG 4335

RESULT 10  
AAA62010  
ID AAA62010 standard; DNA; 4473 BP.  
XX  
XX AAA62010;  
AC  
XX  
XX  
DT 02-FEB-2001 (first entry)  
XX  
XX  
XX Hydrophobic domain protein cDNA HP02837 isolated from HT-1080 cells.  
XX Human; secreted protein; membrane protein; hydrophobic domain;  
KW proliferation control; differentiation induction; material transport;

KM biophylaxis; signal receptor; ion channel; transporter; immunostimulant;  
 KM immunosuppressant; haematopoiesis regulator; chemotactic; chemokinetic;  
 KM haemostatic; thrombolytic; anti-inflammatory; tumour inhibition;  
 KM autoimmune disease; Alzheimer's disease; Parkinson's disease; cancer;  
 KM gene therapy; ss.  
 XX  
 OS Homo sapiens.  
 XX MO200029448-A2.  
 XX  
 PD 25-MAY-2000.  
 XX  
 XX 17-NOV-1999; 99MO-JP06412.  
 XX  
 XX 17-NOV-1998; 98JP-0326255.  
 XX 22-DEC-1998; 98JP-0364315.  
 XX 16-MAR-1999; 99JP-0069811.  
 XX 27-APR-1999; 99JP-0119299.  
 XX 19-MAY-1999; 99JP-0138169.  
 XX  
 PA (SAGA ) SAGAMI CHEM RES CENT.  
 PA (PROT-) PROTEGENE INC.  
 XX  
 PI Kato S, Kimura T;  
 XX  
 DR WPI; 2000-387753/33.  
 DR P-PSDB; AAB12127.  
 XX  
 XX Proteins comprising hydrophobic regions, such as secretory and membrane  
 PT proteins, useful in research and diagnostics and having various  
 PT activities e.g. immunomodulatory, anti-inflammatory, chemokinetic,  
 PT hemostatic, thrombolytic -  
 XX  
 PS Claim 4; Page 253-261; 410pp; English.  
 XX  
 CC Secretory proteins play important roles in the proliferation control, the  
 CC differentiation induction, the material transport and the biophylaxis of  
 CC cells. Membrane proteins have important roles as signal receptors, ion  
 CC channels and transporters. The present sequence is the coding sequence  
 CC for a human protein which has at least one hydrophobic domain. The  
 CC protein encoded by the present sequence may be a secretory or a membrane  
 CC protein. The encoded protein may have cytokine and cell  
 CC proliferation/differentiation activity, immune stimulating or suppressing  
 CC activity, haematopoiesis activity, tissue growth activity,  
 CC activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic  
 CC and thrombolytic activity, anti-inflammatory activity and tumour  
 CC inhibition activity. The present sequence could therefore be used for  
 CC treatment of autoimmune disease, Alzheimer's disease, Parkinson's  
 CC disease, and cancer via gene therapy.  
 CC  
 XX Sequence 4473 BP; 1328 A; 867 C; 952 G; 1326 T; 0 other;  
 SQ  
 Query Match 98.4%; Score 4219.6; DB 21; Length 4473;  
 Bseq Local Similarity 98.7%; Freq. No. 0;  
 Matches 4283; Conservative 0; Mismatches 4; Indels 51; Gaps 1;  
 QY 1 ATGGAGGGCCCAACGGCTCTGACCGCGCCACCTCCCTGCTGTGACCGCGCGCTG 60  
 Db 45 ATGGAGGGCCCAACGGCTCTGACCGCGCCACCTCCCTGCTGTGACCGCGCGCTG 104  
 QY 61 GCCGTGGCTCCCGGGCCCTCGGTTTCTGTGACAGCCCGAGGATCATCAGGCCCGAGGA 120  
 Db 105 GCCGTGGCTCCCGGGCCCTCGGTTTCTGTGACAGCCCGAGGATCATCAGGCCCGAGGA 164  
 QY 121 AATGTGACTATTGGGGGTGAGCTTCTGGAACATGCGCTTCAACAGGTGACTGTGAAGCG 180  
 Db 165 AATGTGACTATTGGGGGTGAGCTTCTGGAACATGCGCTTCAACAGGTGACTGTGAAGCG 224  
 QY 181 GAGGTGCTCAAGACAGCATCAACCTCACTGCTCTGCTCGAGAGCAAGAGAGTCTT 240  
 Db 225 GAGGTGCTCAAGACAGCATCAACCTCACTGCTCTGCTCGAGAGCAAGAGAGTCTT 284  
 QY 241 GAAAAAGGCTCTTTAAGACACTTACTTCCATCACTACTGAAACAGTGCAATGAG 300

Db 285 GAAAAAGGCTCTTTAAGACACTTACTTCCATCACTACTGAAACAGTGCAATGAG 344  
 QY 301 ATTATGAGCTACGTGAACCGGACGTACCCAGATGAGATTTTATCTTAATAGTACC 360  
 Db 345 ATTTATGAGCTACGTGAACCGGACGTACCCAGATGAGATTTTATCTTAATAGTACC 404  
 QY 361 CGCTTATCATTTGAGACCAAGAGATATCTGCTTCAATCAACAGCAAGGCTTATAC 420  
 Db 405 CGCTTATCATTTGAGACCAAGAGATATCTGCTTCAATCAACAGCAAGGCTTATAC 464  
 QY 421 AAGCCAAAGCAAGATGAACTTGGCATTTGTAAGCTCTGCTCAATTTTAAAGCTTAC 480  
 Db 465 AAGCCAAAGCAAGATGAACTTGGCATTTGTAAGCTCTGCTCAATTTTAAAGCTTAC 524  
 QY 481 AAAACCTCTTTAAACATCTCATTTAAGACCCCAATCAAAATTTGATCCACAGTGTG 540  
 Db 525 AAAACCTCTTTAAACATCTCATTTAAGACCCCAATCAAAATTTGATCCACAGTGTG 584  
 QY 541 TCACAACAAAAGTATCTTGAGATCATTTCCAAAACCTTTCAAGTATCTTCCATCAATA 600  
 Db 585 TCACAACAAAAGTATCTTGAGATCATTTCCAAAACCTTTCAAGTATCTTCCATCAATA 644  
 QY 601 CTGTGAGCTGCTCTATTCAGATTCAGATGAAATGACACAGCATATTAATCATTTTAC 660  
 Db 645 CTGTGAGCTGCTCTATTCAGATTCAGATGAAATGACACAGCATATTAATCATTTTAC 704  
 QY 661 GTTTCAGATATGATATACCAAAATTTGAAGTACCTTGACAGACCATTAATGTTCT 720  
 Db 705 GTTTCAGATATGATATACCAAAATTTGAAGTACCTTGACAGACCATTAATGTTCT 764  
 QY 721 ATGAATTTCAAGCATTTAAATGATGATACCATCAGGCAAGATATCATATGAGGAAGCCAGTG 780  
 Db 765 ATGAATTTCAAGCATTTAAATGATGATACCATCAGGCAAGATATCATATGAGGAAGCCAGTG 824  
 QY 781 AAGAGAGCTAACGCTTACATTTTACCTTTATCTTTGGGAAAAGAAATATAT 840  
 Db 825 AAGAGAGCTAACGCTTACATTTTACCTTTATCTTTGGGAAAAGAAATATAT 884  
 QY 841 ACAAACAATTTAAGATTAATGATGATGCAAACTTCTTTTAAATGATGAAGATGA 900  
 Db 885 ACAAACAATTTAAGATTAATGATGATGCAAACTTCTTTTAAATGATGAAGATGA 944  
 QY 901 AATGATATGATTTCTTCAATGAGATCTTGTGATACCTGATCATCTTCCCTGAGCA 960  
 Db 945 AATGATATGATTTCTTCAATGAGATCTTGTGATACCTGATCATCTTCCCTGAGCA 1004  
 QY 961 GTTGAATTTTAAACCAAGTACAGATGATGATGATGATTTCAAGAAATGTAAGCACT 1020  
 Db 1005 GTTGAATTTTAAACCAAGTACAGATGATGATGATGATTTCAAGAAATGTAAGCACT 1064  
 QY 1021 AATGTGCTCTCAAGACATGATTTACATCATTTGATTTTGAATATCTACTGCTTG 1080  
 Db 1065 AATGTGCTCTCAAGACATGATTTACATCATTTGATTTTGAATATCTACTGCTTG 1124  
 QY 1081 AAGCCATCTCTCAACTTCAAGCAAGCACTGTGAAGTAACTCGTGTGATGAGCAACAATG 1140  
 Db 1125 AAGCCATCTCTCAACTTCAAGCAAGCACTGTGAAGTAACTCGTGTGATGAGCAACAATG 1184  
 QY 1141 ACTCTTGAAGAAAGAAATTAATGATGATTAATGATGATGATGATGATGATGATG 1200  
 Db 1185 ACTCTTGAAGAAAGAAATTAATGATGATTAATGATGATGATGATGATGATGATG 1244  
 QY 1201 TACTGAGAGGAGATCTTAACAGTGAATTCAGAAATGGAAGCTGTTCAGAAATTAATAT 1260  
 Db 1245 TACTGAGAGGAGATCTTAACAGTGAATTCAGAAATGGAAGCTGTTCAGAAATTAATAT 1304  
 QY 1261 ACTGTCCCCCAAGTGAACCTTTTAAGATTTGAATTTCCATCTCTGAGAGATTCAGTGA 1320  
 Db 1305 ACTGTCCCCCAAGTGAACCTTTTAAGATTTGAATTTCCATCTCTGAGAGATTCAGTGA 1364  
 QY 1321 CTACAGTTAAGGCTTATTTCTTGTGATGAAATGATGATGATGATGATGATGATG 1380



Db 1365 CTACAGTTGAAGCCCTATTTCTTGGTAGTAAAAAGACATGGCAGTTCAATGCTGTTT 1424  
Qy 1381 AAGTCTCTAGTAAGACATACATCCAACTAAAAACAAGATGAATAATATAAGTGGGA 1440  
Db 1425 AAGTCTCTAGTAAGACATACATCCAACTAAAAACAAGATGAATAATATAAGTGGGA 1484  
Qy 1441 TGGCCTTTTGAAGTGGTGGTGAAGTGGCAAAACGATGAGAGGATTAAGCTAATATGTA 1500  
Db 1485 TGGCCTTTTGAAGTGGTGGTGAAGTGGCAAAACGATGAGAGGATTAAGCTAATATGTA 1544  
Qy 1501 GATATCCAGGGGACAGTTGGTGGCTGAAGAAAAAAATTCACACATGTTCTCTTTAA 1560  
Db 1545 GATATCCAGGGGACAGTTGGTGGCTGAAGAAAAAAATTCACACATGTTCTCTTTAA 1604  
Qy 1561 CCAGAAAAATTTCTGACCTCAAAAGCCTGTGTAATGTGATTAATTAAGATGATG 1620  
Db 1605 CCAGAAAAATTTCTGACCTCAAAAGCCTGTGTAATGTGATTAATTAAGATGATG 1664  
Qy 1621 GAAATTAATAAGATGATGTTCTAAATAATCTGTTCACCTGTTTTTAAAAATAGATAAG 1680  
Db 1665 GAAATTAATAAGATGATGTTCTAAATAATCTGTTCACCTGTTTTTAAAAATAGATAAG 1724  
Qy 1681 CTATATGAGTAAGTAAGTAAGTGAACCATCTGAGAAAAGTCTCTTAGATCTCTGTG 1740  
Db 1725 CTATATGAGTAAGTAAGTAAGTGAACCATCTGAGAAAAGTCTCTTAGATCTCTGTG 1784  
Qy 1741 ACACAGCTGACCTCATAGTTGGGATGTAGCTGTGACAAAAGTGAATCTGATGAAT 1800  
Db 1785 ACACAGCTGACCTCATAGTTGGGATGTAGCTGTGACAAAAGTGAATCTGATGAAT 1844  
Qy 1801 GCCTCTAATGATATTAACAATGAAAAATGTGTCCATGAGTTGGAACTTTATTAACACAGA 1860  
Db 1845 GCCTCTAATGATATTAACAATGAAAAATGTGTCCATGAGTTGGAACTTTATTAACACAGA 1904  
Qy 1861 TATATATTTAGGCATGTTCAATGAAATCTTTTGCACTCTTTCAGGAATGTGACTGTGGTA 1920  
Db 1905 TATATATTTAGGCATGTTCAATGAAATCTTTTGCACTCTTTCAGGAATGTGACTGTGGTA 1964  
Qy 1921 TTGACAGATGCAAACTCGAAGAGTATTAATGATGTTTANGACAATGACAGAAAT 1980  
Db 1965 TTGACAGATGCAAACTCGAAGAGTATTAATGATGTTTANGACAATGACAGAAAT 2024  
Qy 1981 GCTGAGAGTATTAAGAGAAAAATGAAGACATATGTAGATTAATCATGACTTTCTTGG 2040  
Db 2025 GCTGAGAGTATTAAGAGAAAAATGAAGACATATGTAGATTAATCATGACTTTCTTGG 2084  
Qy 2041 GGTAGAGTCCACATGTCCGAAAGCAATTTCCAGACATTTGAGTGTGAGTACACCAAC 2100  
Db 2085 GGTAGAGTCCACATGTCCGAAAGCAATTTCCAGACATTTGAGTGTGAGTACACCAAC 2144  
Qy 2101 ATGGGTTACAGATTTTACCAAGAAATTTGAAGTAAGTAAGTCTGATCTATCACTTTCTGG 2160  
Db 2145 ATGGGTTACAGATTTTACCAAGAAATTTGAAGTAAGTAAGTCTGATCTATCACTTTCTGG 2204  
Qy 2161 GTGGCTACTGTTTGTGATCTCTGAGACCTGGGCTTGGACTGACAACTACTCCAGTG 2220  
Db 2205 GTGGCTACTGTTTGTGATCTCTGAGACCTGGGCTTGGACTGACAACTACTCCAGTG 2264  
Qy 2221 GAGCTCCAAAGCCTTCCAAACATTTTCAATTTTTTGAATCTTCCCTACTCTGTTATCAGA 2280  
Db 2265 GAGCTCCAAAGCCTTCCAAACATTTTCAATTTTTTGAATCTTCCCTACTCTGTTATCAGA 2324  
Qy 2281 GGTGAAGATTTTCTTGGAAAAATTAATATTAATTAATTTTGAAGAATGCCCTGAGGTT 2340  
Db 2325 GGTGAAGATTTTCTTGGAAAAATTAATATTAATTAATTTTGAAGAATGCCCTGAGGTT 2384  
Qy 2341 AAGGTATCATTTGAGAAAAAGTGAACAATTTGATATTCTAATGACTTCAAGTGAATTAAT 2400  
Db 2385 AAGGTATCATTTGAGAAAAAGTGAACAATTTGATATTCTAATGACTTCAAGTGAATTAAT 2444  
Qy 2401 GCCACAGGCCACACAGACACCTTGTGTTCCAGTGAAGATGGGCAACTGTTCTTTT 2460  
Db 2445 GCCACAGGCCACACAGACACCTTGTGTTCCAGTGAAGATGGGCAACTGTTCTTTT 2504  
Qy 2461 CCATCAGGCCACACATCTGGAGAAATTCCTATCAGATCAAGCTCTTTCACCCACT 2520  
Db 2505 CCATCAGGCCACACATCTGGAGAAATTCCTATCAGATCAAGCTCTTTCACCCACT 2564  
Qy 2521 GCTTCTGATGCTGTGACCCCAAGATGATTTTGAAGAGCTGAAGAAATGAATAATCATAT 2580  
Db 2565 GCTTCTGATGCTGTGACCCCAAGATGATTTTGAAGAGCTGAAGAAATGAATAATCATAT 2624  
Qy 2581 TCACAAATCCATTTTATTTGACTTGAATGACAAATAGGCTTACAGAGTACCCCTGAAAACTTTG 2640  
Db 2625 TCACAAATCCATTTTATTTGACTTGAATGACAAATAGGCTTACAGAGTACCCCTGAAAACTTTG 2684  
Qy 2641 AGTTCTCATTTTCCCTCATATACAGTGAATGAGTGAAGAAAGTTCAGATCAATGCAAT 2700  
Db 2685 AGTTCTCATTTTCCCTCATATACAGTGAATGAGTGAAGAAAGTTCAGATCAATGCAAT 2744  
Qy 2701 GAGAGATGTTCTTGGTCTTTCATCAATGAGCTTGAAGCTTGAATTCAGATGCTTATGAGC 2760  
Db 2745 GAGAGATGTTCTTGGTCTTTCATCAATGAGCTTGAAGCTTGAATTCAGATGCTTATGAGC 2804  
Qy 2761 TGTGTGAACAGACATGATTAATTTGCTGCAATATTTTCAATTTTGGATTAATCTGACT 2820  
Db 2805 TGTGTGAACAGACATGATTAATTTGCTGCAATATTTTCAATTTTGGATTAATCTGACT 2864  
Qy 2821 AAAAAAACAACAGACATTAATTTGAAGAAAAAGCTCTTTCATTTATGAGGCAAGT 2880  
Db 2865 AAAAAAACAACAGACATTAATTTGAAGAAAAAGCTCTTTCATTTATGAGGCAAGT 2924  
Qy 2881 TACAGAGAGAACTTCTATCAGAGGAGATGAGCTCTTTCAGTCTTTTGGAAATTA 2940  
Db 2925 TACAGAGAGAACTTCTATCAGAGGAGATGAGCTCTTTCAGTCTTTTGGAAATTA 2984  
Qy 2941 GACCTTCTGGAGAGACTTGGTGTGACCTTTTGTTTTAAAGTGTTCCTTGAAGCCGAT 3000  
Db 2985 GACCTTCTGGAGAGACTTGGTGTGACCTTTTGTTTTAAAGTGTTCCTTGAAGCCGAT 3044  
Qy 3001 CCTTACATGATTAATGATCAGATGTGTGTAACAAGAAATCACTTGGCTTAAAGGAAT 3060  
Db 3045 CCTTACATGATTAATGATCAGATGTGTGTAACAAGAAATCACTTGGCTTAAAGGAAT 3104  
Qy 3061 CAGAAATCCAAAGGTAATTTTGGATCAGAGAGAGATTCATATGAGACTTCAAGAT 3120  
Db 3105 CAGAAATCCAAAGGTAATTTTGGATCAGAGAGAGATTCATATGAGACTTCAAGAT 3164  
Qy 3121 GGCATTAATAAGTCCAGTAACATTTACAGCTTATATTTGTAATCTTCTCTGGGATATAGA 3180  
Db 3165 GGCATTAATAAGTCCAGTAACATTTACAGCTTATATTTGTAATCTTCTCTGGGATATAGA 3224  
Qy 3181 AAGTATCAGCCCTAATGATGATGTCAGAGAGTCAATCCATTTTGGAGTCTGAATTCAGT 3240  
Db 3225 AAGTATCAGCCCTAATGATGATGTCAGAGAGTCAATCCATTTTGGAGTCTGAATTCAGT 3284  
Qy 3241 AGAGGAATTTCAACAATTAATCTAGCCCTTATTAATTTTANGACTTTCATCAGTGGGG 3300  
Db 3285 AGAGGAATTTCAACAATTAATCTAGCCCTTATTAATTTTANGACTTTCATCAGTGGGG 3344  
Qy 3301 AGTCTTAAGCGAAGAGCTTTGAAATGTCTGACTTGAAGAGCAAAAGAGTGGC 3360  
Db 3345 AGTCTTAAGCGAAGAGCTTTGAAATGTCTGACTTGAAGAGCAAAAGAGTGGC 3404  
Qy 3361 ATGCAATCTGGGTGTCTACAGAGTCCAAATTTTCTGACTCTGGACCCAGCTCCCTG 3420  
Db 3405 ATGCAATCTGGGTGTCTACAGAGTCCAAATTTTCTGACTCTGGACCCAGCTCCCTG 3464  
Qy 3421 GATATTTGAAGTTGACGCTATGACCTGCTCACAATTTTCAATTTTCAAGCTTCTGAG 3480  
Db 3465 GATATTTGAAGTTGACGCTATGACCTGCTCACAATTTTCAAGCTTCTTCAAGCTTCTGAG 3524  
Qy 3481 GGAATCCAAATTTATGAGTGTGCTTAAGCAGGCAAGAAATAGCTTGGTGTGTTGATCT 3540  
Db 3525 GGAATCCAAATTTATGAGTGTGCTTAAGCAGGCAAGAAATAGCTTGGTGTGTTGATCT 3584

QY	3541	ACTAGAGTATACACAGTGGCTTTAAAGGCTGTCGTGAATTTGGAGCCCTPATGAATACA	3600
Db	3585	ACTCAGGATACCACTGAGCTTTAAAGGCTCTGTCTGAATTTGGAGCCCTPATGAATACA	3644
QY	3601	GAAGAACAATATCCAGAGTACCGCTGACGGGGCTTAGCTACCAAGTCTCT-----	36511
Db	3645	GAAGAACAATATCCAGAGTACCGCTGACGGGGCTTAGCTACCAAGTCTCTGAAGTTT	3704
QY	3652	-----CTTGGCTGTGGTACAGCCA	3669
Db	3705	CTGATTGCACACACAAACCGCTTACCTCTTGACAGCAGAGCTTGGCTGTGGTACAGCCA	3764
QY	3670	ATGGCAGTTAATATATTCGGCAAAATGGTTTGGATTTGGTATTTGGCAGCTCAATGGTTGA	3729
Db	3765	ACGGCAGTTAATATATTCGGCAAAATGGTTTGGATTTGGTATTTGGCAGCTCAATGGTTGA	3824
QY	3730	TATATATGAAAGGCTTCTGGGTCTTCTGAAAGACGAGATCTATCCAAATCAAGAAACC	3789
Db	3825	TATATATGAAAGGCTTCTGGGTCTTCTGAAAGACGAGATCTATCCAAATCAAGAAACC	3884
QY	3790	TTTGATTTAGATGTTGCTGTGTAAGAAAAATTAAGATGATCTCATCATGTGGATTTGAT	3849
Db	3885	TTTGATTTAGATGTTGCTGTGTAAGAAAAATTAAGATGATCTCATCATGTGGATTTGAT	3944
QY	3850	GTGTGTACAAAGCTTTGGGGCCCGGGGTGAGAGTGCGAGCGCTTATGGAAGTTAACTTA	3909
Db	3945	GTGTGTACAAAGCTTTGGGGCCCGGGGTGAGAGTGCGAGCGCTTATGGAAGTTAACTTA	4004
QY	3910	TTAAGTGGCTTATGAGTGCCTTCTGAGAAACAATTTCTGAGCGAGACAGTAAAGAAATG	3969
Db	4005	TTAAGTGGCTTATGAGTGCCTTCTGAGAAACAATTTCTGAGCGAGACAGTAAAGAAATG	4064
QY	3970	GAATATGATCAATGAAAACTCAACCTCTATTAGATTCTGTAATGAAACCCAGTTTGT	4029
Db	4065	GAATATGATCAATGAAAACTCAACCTCTATTAGATTCTGTAATGAAACCCAGTTTGT	4124
QY	4030	GTTAATATTCCTGCTGTGAGAAACCTTTAAAGTTCAATTAACCAAGATGCTTCAGTGTCC	4089
Db	4125	GTTAATATTCCTGCTGTGAGAAACCTTTAAAGTTCAATTAACCAAGATGCTTCAGTGTCC	4184
QY	4090	ATATGATGATTACTATGAGCCAAAGAGACAGCGCGTGAAGATTCAACTCTGAAGTGAAG	4149
Db	4185	ATATGATGATTACTATGAGCCAAAGAGACAGCGCGTGAAGATTCAACTCTGAAGTGAAG	4244
QY	4150	CTGTCTCTCTGTGACCTTTGACAGTATGTCCAGGGCTGCGGTCTTGTGAGAAATGAGACT	4209
Db	4245	CTGTCTCTCTGTGACCTTTGACAGTATGTCCAGGGCTGCGGTCTTGTGAGAAATGAGACT	4304
QY	4210	TCAGAGCTCCCAATCAATCACTCTTCAGTCAATTTTATTTTCTGTTCAGACTCTGTACTTT	4289
Db	4305	TCAGAGCTCCCAATCAATCACTCTTCAGTCAATTTTATTTTCTGTTCAGACTCTGTACTTT	4364
QY	4270	ATGGAACCTTGGAGTGTGA 4287	
Db	4365	ATGGAACCTTGGAGTGTGA 4382	
RESULT 11			
AAA60199			
ID	AAA60199	standard; DNA; 4335 BP.	
XX	AAA60199;		
DT	02-FEB-2001	(first entry)	
DE	Hydrophobic domain protein cDNA HP02837 isolated from HT-1080 cells.		
KX	Human; secreted protein; membrane protein; hydrophobic domain;		
KM	proliferation control; differentiation induction; material transport;		
KM	biophysics; signal receptor; ion channel; transporter; immunostimulant;		
KM	immunosuppressant; haematopoiesis regulator; chemotactic; chemokinetic;		
KM	haemostatic; thrombolytic; anti-inflammatory; tumour inhibition;		
KM	autoimmune disease; Alzheimer's disease; Parkinson's disease; cancer;		

Query Match	98.4%	Score 4216.6	DB 21	Length 4335
Beet Local Similarity	98.7%	Pred. No. 0		
Matches 4280	Conservative 0	Mismatches 4	Indels 51	Gaps 1
1 ATGCAGGAGCCACCGCTCTGTACCGCGCCACCTCTCTGTCGTGACCGCGCGCTG	60			
1 ATGCAGGAGCCACCGCTCTGTACCGCGCCACCTCTCTGTCGTGACCGCGCGCTG	60			
61 GCCGTGCTCCCGGCTCGGTTTCTGTGACAGCCCCAGGATCATCGCCCGGAGGA	120			
61 GCCGTGCTCCCGGCTCGGTTTCTGTGACAGCCCCAGGATCATCGCCCGGAGGA	120			
121 AATGTGACTATTGGGGGTGAGCTTCTGGAACACCTGCCCTTCAACAGGTGACTGTGAAGCG	180			
121 AATGTGACTATTGGGGGTGAGCTTCTGGAACACCTGCCCTTCAACAGGTGACTGTGAAGCG	180			
181 GAGGTGCTCAAGACAGATCAAACTGACGTCTGTGCTCGGAGGAGGAAGAGTCTTT	240			
181 GAGGTGCTCAAGACAGATCAAACTGACGTCTGTGCTCGGAGGAGGAAGAGTCTTT	240			
241 GAAAAAGGCTCTTTTAAGACACTTACTCTTCCATCACTACTCTGAAACAGTGACAGATGAG	300			
241 GAAAAAGGCTCTTTTAAGACACTTACTCTTCCATCACTACTCTGAAACAGTGACAGATGAG	300			
301 ATTATGAGCTACGTGTAACCGGACGTAACCCAGATGAGATTTTATTCTTAATGATACC	360			

```
Db 301 ATTTAGAGCTAGCTGTAACCGGACGTACCAGAGATGAGATTTATCTTAATAGTACC 360
Qy 361 CCCTTATCATTTGAGACCAAGAAATATCTGCTTCAATCAACAGACAGAGGCTTAATAC 420
Db 361 CGCTTATCATTTGAGACCAAGAAATATCTGCTTCAATCAACAGACAGAGGCTTAATAC 420
Qy 421 AAGCCAAAGCAAGAAAGTGAAGTTGCGCATGTTTACATCTCTTCAGATTTTAAAGCTTAC 480
Db 421 AAGCCAAAGCAAGAAAGTGAAGTTGCGCATGTTTACATCTCTTCAGATTTTAAAGCTTAC 480
Qy 481 AAAAACCCTTTTAAACATTTCTCATTTAAGAACCCCAATCAAAATTTGATCCAAAGTGGTG 540
Db 481 AAAAACCCTTTTAAACATTTCTCATTTAAGAACCCCAATCAAAATTTGATCCAAAGTGGTG 540
Qy 541 TCACAAACAAGTGAATCTGAGATCATTTCCAAAACCTTTCAAGTATCTTCCCATCCAAAT 600
Db 541 TCACAAACAAGTGAATCTTGGAGTCATTTCCAAAACCTTTCAAGTATCTTCCCATCCAAAT 600
Qy 601 CTGTGAGCTGCTATTCAGTCAAGTTCAAGTGAACAGACATATTAATCAATCATTTTACAG 660
Db 601 CTGTGAGCTGCGTCTATTCAGTCAAGTGAAGTGAACAGACATATTAATCAATCATTTTACAG 660
Qy 661 GTTTTCGAAATATGATTTACCAAAATTTGAAGTGAATTTGCGAAGACCATTAATATGTTCT 720
Db 661 GTTTTCGAAATATGATTTACCAAAATTTGAAGTGAATTTGCGAAGACCATTAATATGTTCT 720
Qy 721 ATGAATTTCAAGATTTAAATGATGACATACAGGCAAAATATATCAATATGGAAGCCAGAG 780
Db 721 ATGAATTTCAAGATTTAAATGATGACATACAGGCAAAATATATCAATATGGAAGCCAGAG 780
Qy 781 AAAGAGACGTAAACGTTCATTTTACCTTTATCTTTTGGGGAAGAAAGAAATAT 840
Db 781 AAAGAGACGTAAACGTTCATTTTACCTTTATCTTTTGGGGAAGAAAGAAATAT 840
Qy 841 ACAAACCAATTTAAGTAAATGAGATCTGCAAACTTCTCTTTTATGATGAAGATGAAA 900
Db 841 ACAAACCAATTTAAGTAAATGAGATCTGCAAACTTCTCTTTTATGATGAAGATGAAA 900
Qy 901 AATGTAAATGATTTCTTCAATGGAATCTTCTGAAATACCTGATCTATCTTCCCTGAGACA 960
Db 901 AATGTAAATGATTTCTTCAATGGAATCTTCTGAAATACCTGATCTATCTTCCCTGAGACA 960
Qy 961 GTAGAAATTTTAAACCAAGTGAACAGATCAGATTAAGATTTTCAAGAAATGTAAGCACT 1020
Db 961 GTAGAAATTTTAAACCAAGTGAACAGATCAGATTAAGATTTTCAAGAAATGTAAGCACT 1020
Qy 1021 AATGTGTTCTTCAAGCAATGATTAATCATTTGAGTTTGTATTAATCTATCTGTCTTG 1080
Db 1021 AATGTGTTCTTCAAGCAATGATTAATCATTTGAGTTTGTATTAATCTATCTGTCTTG 1080
Qy 1081 AAGCCATCTCTCAATCTTCAAGCACTGGAAGTAACTGTCGTAAGTGAACCAACCTG 1140
Db 1081 AAGCCATCTCTCAATCTTCAAGCACTGGAAGTAACTGTCGTAAGTGAACCAACCTG 1140
Qy 1081 AAGCCATCTCTCAATCTTCAAGCACTGGAAGTAACTGTCGTAAGTGAACCAACCTG 1140
Db 1081 AAGCCATCTCTCAATCTTCAAGCACTGGAAGTAACTGTCGTAAGTGAACCAACCTG 1140
Qy 1141 ACTCTTGAAGAAAGAAATATATGATCATPAACAGTGAACAGAGAACTATATCTGAG 1200
Db 1141 ACTCTTGAAGAAAGAAATATATGATCATPAACAGTGAACAGAGAACTATATCTGAG 1200
Qy 1201 TACTGAGCGGATCTTAACAGTGAATTCAGAAAATGGAAGCTGTTTCAAGAAATTAAT 1260
Db 1201 TACTGAGCGGATCTTAACAGTGAATTCAGAAAATGGAAGCTGTTTCAAGAAATTAAT 1260
Qy 1261 ACTGTCCCCCAAGTGAATCTTTAAGTGAATTCCTCAATCTGAGGATTCAGTGAAG 1320
Db 1261 ACTGTCCCCCAAGTGAATCTTTAAGTGAATTCCTCAATCTGAGGATTCAGTGAAG 1320
Qy 1321 CTACAGTTGAAGGCTATTTCTTGTAGTAAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1380
Db 1321 CTACAGTTGAAGGCTATTTCTTGTAGTAAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1380
Qy 1381 AAGTCTCTAGTAAAGCATATCATCAATTAACAAAGATGAAGATTAAGTGAAGTGAAG 1440
Db 1381 AAGTCTCTAGTAAAGCATATCATCAATTAACAAAGATGAAGATTAAGTGAAGTGAAGTGAAG 1440
```

---

```
Db 1381 AAGTCTCTAGTAAAGCATATCATCAATTAACAAAGATGAAGATTAAGTGAAGTGAAG 1440
Qy 1441 TCGCCTTTTGAAGTGGTGGTATGAGCAACAAAGATGAAGAGTGAAGTATATGAGTA 1500
Db 1441 TCGCCTTTTGAAGTGGTGGTATGAGCAACAAAGATGAAGAGTGAAGTATATGAGTA 1500
Qy 1501 GTATCAGAGGAGCAGTTGGTGGTGGTGAAGAAACAAATTCACACATGTTCTTTTAAACA 1560
Db 1501 GTATCAGAGGAGCAGTTGGTGGTGGTGAAGAAACAAATTCACACATGTTCTTTTAAACA 1560
Qy 1561 CCAAGAAATTTTGGAGCTCCAAAAGCTGTGTAAATTTGATATTAATGAAGATGATGGG 1620
Db 1561 CCAAGAAATTTTGGAGCTCCAAAAGCTGTGTAAATTTGATATTAATGAAGATGATGGG 1620
Qy 1621 GAATTTATAGTGAATGTTCTTAAATTTCTGTTCAGTGTGTTTAAATTAAGTAAAG 1680
Db 1621 GAATTTATAGTGAATGTTCTTAAATTTCTGTTCAGTGTGTTTAAATTAAGTAAAG 1680
Qy 1681 CTATATTTGAGTGAAGTGAAGAGTGAACCATCTGAGAAAGTCTCTTGAAGATCTGTG 1740
Db 1681 CTATATTTGAGTGAAGTGAAGAGTGAACCATCTGAGAAAGTCTCTTGAAGATCTGTG 1740
Qy 1741 ACAAGCCTGACTCCATATGTTGGATTTGAGCTGTGAACAAAGTGAATCTGATGAAT 1800
Db 1741 ACAAGCCTGACTCCATATGTTGGATTTGAGCTGTGAACAAAGTGAATCTGATGAAT 1800
Qy 1801 GCTCTAATGATTAATTAAGTGAAGAAATGCTGCAATGAGTGAACCTTATTAACAGAGA 1860
Db 1801 GCTCTAATGATTAATTAAGTGAAGAAATGCTGCAATGAGTGAACCTTATTAACAGAGA 1860
Qy 1861 TATTAATTTAGCAGATGTTCAATGAATCTTTTCAAGCTTTTCAAGAAATGAGACTGAGTA 1920
Db 1861 TATTAATTTAGCAGATGTTCAATGAATCTTTTCAAGCTTTTCAAGAAATGAGACTGAGTA 1920
Qy 1921 TTGACAGATGCAAACTTCAAGAAAGTGAATTAATGATGTTTATGAACAAAGCAGAAATAT 1980
Db 1921 TTGACAGATGCAAACTTCAAGAAAGTGAATTAATGATGTTTATGAACAAAGCAGAAATAT 1980
Qy 1981 GCTGAGAGTGAATGAAGAAATGAAGAAATGAAGAAATGAAGAAATGAAGAAATGAAGAAAT 2040
Db 1981 GCTGAGAGTGAATGAAGAAATGAAGAAATGAAGAAATGAAGAAATGAAGAAATGAAGAAAT 2040
Qy 2041 GGTAGCAGTCCACATGTCGGAAGAAAGATTTTCCAGAGACTTGGATTTGGCTAGACCAAC 2100
Db 2041 GGTAGCAGTCCACATGTCGGAAGAAAGATTTTCCAGAGACTTGGATTTGGCTAGACCAAC 2100
Qy 2101 ATGGGTTACAGGATTTTACCAAGAAATTTGAAGTAACTGATCTGATCACTTCTTGG 2160
Db 2101 ATGGGTTACAGGATTTTACCAAGAAATTTGAAGTAACTGATCTGATCACTTCTTGG 2160
Qy 2161 GTGGCTACTGGTTTGTGATCTGAGAGACCTGGGCTTGGACATTAACATCTCCAGATG 2220
Db 2161 GTGGCTACTGGTTTGTGATCTGAGAGACCTGGGCTTGGACATTAACATCTCCAGATG 2220
Qy 2221 GAGCTCCAAAGCTTCCAAACATTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 2280
Db 2221 GAGCTCCAAAGCTTCCAAACATTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 2280
Qy 2281 GGTGAAGATTTGCTTTTGAATTAATATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2340
Db 2281 GGTGAAGATTTGCTTTTGAATTAATATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2340
Qy 2341 AAGGTAAATCAATTTGAAGAAAGTGAACAAATTTGAATTTGAAGAAATGAAGAAATTAAT 2400
Db 2341 AAGGTAAATCAATTTGAAGAAAGTGAACAAATTTGAATTTGAAGAAATGAAGAAATTAAT 2400
Qy 2401 GCCACAGGCAACAGACAGACCTTCTGGTTCAGGTGAGATGAGGAGCACTGTTCTTTT 2460
Db 2401 GCCACAGGCAACAGACAGACCTTCTGGTTCAGGTGAGATGAGGAGCACTGTTCTTTT 2460
Qy 2461 CCCATCAGGCAACAGATCTGGAGAAATTTCTATCAAGTCAAGCTCTTTTCAACCACT 2520
Db 2461 CCCATCAGGCAACAGATCTGGAGAAATTTCTATCAAGTCAAGCTCTTTTCAACCACT 2520
```

QY 2521 GCTTCTGATGCTGTCACCCAGATGATTTTAAAGGCTGAAGAAATAGAAAAATCATAT 2580  
 Db 2521 GCTTCTGATGCTGTCACCCAGATGATTTTAAAGGCTGAAGAAATAGAAAAATCATAT 2580  
 QY 2581 TCACATTCATCTTATTTAGACTGACACATAGGCTGACAGAGTACCTGAAAATTG 2640  
 Db 2581 TCACATTCATCTTATTTAGACTGACACATAGGCTGACAGAGTACCTGAAAATTG 2640  
 QY 2641 AGTTCTCATCTTCCCTTAATACAGTGA CTGGCAGTGAAGAGTTCAGATACAGCAATT 2700  
 Db 2641 AGTTCTCATCTTCCCTTAATACAGTGA CTGGCAGTGAAGAGTTCAGATACAGCAATT 2700  
 QY 2701 GGAGATGTTCTTGGCTCTTCATGCAATGCTTAAGCTTCATGATTCGGATCCCTTAAGC 2760  
 Db 2701 GGAGATGTTCTTGGCTCTTCATGCAATGCTTAAGCTTCATGATTCGGATCCCTTAAGC 2760  
 QY 2761 TGTGCTGAACAGAACATGATTAATTTGCTCCAAATATTACATTTTGATTAATCTGACT 2820  
 Db 2761 TGTGCTGAACAGAACATGATTAATTTGCTCCAAATATTACATTTTGATTAATCTGACT 2820  
 QY 2821 AAAAAGAAACAACAGACATTAATTTGAAGAAAAGCTTTGATTTATGAGGCAAGT 2880  
 Db 2821 AAAAAGAAACAACAGACATTAATTTGAAGAAAAGCTTTGATTTATGAGGCAAGT 2880  
 QY 2881 TACCAAGAGAGACTTCTCTATCAGAGGAGAGATGCTCTTCACTGCTTTTGAGAAATAT 2940  
 Db 2881 TACCAAGAGAGACTTCTCTATCAGAGGAGAGATGCTCTTCACTGCTTTTGAGAAATAT 2940  
 QY 2941 GACCCCTTCTGGAGACCTTGGTGTCACTTTTGTGTTTAAAGATGTTCTTGAAGCCGAT 3000  
 Db 2941 GACCCCTTCTGGAGACCTTGGTGTCACTTTTGTGTTTAAAGATGTTCTTGAAGCCGAT 3000  
 QY 3001 CCTTACATAGATATTGATCAGAAATGTTTACAGAAACATACATGCTGCTTAAAGGACAT 3060  
 Db 3001 CCTTACATAGATATTGATCAGAAATGTTTACAGAAACATACATGCTGCTTAAAGGACAT 3060  
 QY 3061 CAGAAATCCAAAGGCTGATTTTGGGATCCAGAAAGATGATTCATAGAGCTTCAAGGT 3120  
 Db 3061 CAGAAATCCAAAGGCTGATTTTGGGATCCAGAAAGATGATTCATAGAGCTTCAAGGT 3120  
 QY 3121 GGCATTAATAAGCTCAGTAACTTACAGCTAATTTGTAATCTTCTCTGGGATATAGA 3180  
 Db 3121 GGCATTAATAAGCTCAGTAACTTACAGCTAATTTGTAATCTTCTCTGGGATATAGA 3180  
 QY 3181 AAGTATCAGCTTACATTTGATGTCAGAGAGTCAATTTTGGAGTCTGAATTCAGT 3240  
 Db 3181 AAGTATCAGCTTACATTTGATGTCAGAGAGTCAATTTTGGAGTCTGAATTCAGT 3240  
 QY 3241 AGAGGAATTTTCAACAATTTACTCTAGCCCTTAATCTTATGATTTGAGTCTGAGTGGG 3300  
 Db 3241 AGAGGAATTTTCAACAATTTACTCTAGCCCTTAATCTTATGATTTGAGTCTGAGTGGG 3300  
 QY 3301 AGTCTTAAAGGGAAGAGCTTTTGAATATGCTGA CTGGAGAGCAAAAGAGTGGC 3360  
 Db 3301 AGTCTTAAAGGGAAGAGCTTTTGAATATGCTGA CTGGAGAGCAAAAGAGTGGC 3360  
 QY 3361 ATGCAATTCGGGTGATCAGAGTCCAAATTTTCTGATCTCTGAGCCACGCTCCCTG 3420  
 Db 3361 ATGCAATTCGGGTGATCAGAGTCCAAATTTTCTGATCTCTGAGCCACGCTCCCTG 3420  
 QY 3421 GATATTTGAAGTTGAGCTATGCACTGCTCACAATCTTACAAATTTTCAAGCTTCTGAG 3480  
 Db 3421 GATATTTGAAGTTGAGCTATGCACTGCTCACAATCTTACAAATTTTCAAGCTTCTGAG 3480  
 QY 3481 GGAATCCCAATTTATGAGGTGCTAAGCAGGCAAGAAATAGCTTGGTGTGTTGCAATCT 3540  
 Db 3481 GGAATCCCAATTTATGAGGTGCTAAGCAGGCAAGAAATAGCTTGGTGTGTTGCAATCT 3540  
 QY 3541 ACTCAGATACCACTGCTGCTTTAAAGGCTGTGCTGAATTTGCAAGCCCTTAATGAATACA 3600  
 Db 3541 ACTCAGATACCACTGCTGCTTTAAAGGCTGTGCTGAATTTGCAAGCCCTTAATGAATACA 3600

QY 3601 GAAAGACAAATATCCAGTGAACCTGACGGGCTAGCTCAACAAATCTT----- 3651  
 Db 3601 GAAAGACAAATATCCAGTGAACCTGACGGGCTAGCTCAACAAATCTT----- 3651  
 QY 3652 -----CTTGCCTGTGACAGCCA 3669  
 Db 3652 -----CTTGCCTGTGACAGCCA 3669  
 QY 3661 CTGATTGACACACACACACCGCTTACTCTTCAGACAGCAGAGCTTCTGTGTACGCCA 3720  
 Db 3661 CTGATTGACACACACACACCGCTTACTCTTCAGACAGCAGAGCTTCTGTGTACGCCA 3720  
 QY 3670 ATGGAGTTAATATTTTCCGCAATTTGTTTGAATTTGATTTTGTAGCTCAATGTTGTA 3729  
 Db 3670 ATGGAGTTAATATTTTCCGCAATTTGTTTGAATTTGATTTTGTAGCTCAATGTTGTA 3729  
 QY 3721 ACGGCATTAATATTTTCCGCAATTTGTTTGAATTTGATTTTGTAGCTCAATGTTGTA 3780  
 Db 3721 ACGGCATTAATATTTTCCGCAATTTGTTTGAATTTGATTTTGTAGCTCAATGTTGTA 3780  
 QY 3730 TATTAATGTAAGGCTTCTGGCTTCTTAAGAAGCGAAGATCTATCCAAATCAAGAGCC 3789  
 Db 3730 TATTAATGTAAGGCTTCTGGCTTCTTAAGAAGCGAAGATCTATCCAAATCAAGAGCC 3789  
 QY 3781 TATTAATGTAAGGCTTCTGGCTTCTTAAGAAGCGAAGATCTATCCAAATCAAGAGCC 3840  
 Db 3781 TATTAATGTAAGGCTTCTGGCTTCTTAAGAAGCGAAGATCTATCCAAATCAAGAGCC 3840  
 QY 3790 TTTGATTTTGAATTTTGTGCTGTTAAAGAAATTAAGATCTCAATGATGCAATTTGAAT 3849  
 Db 3790 TTTGATTTTGAATTTTGTGCTGTTAAAGAAATTAAGATCTCAATGATGCAATTTGAAT 3849  
 QY 3841 TTTGATTTTGAATTTTGTGCTGTTAAAGAAATTAAGATCTCAATGATGCAATTTGAAT 3900  
 Db 3841 TTTGATTTTGAATTTTGTGCTGTTAAAGAAATTAAGATCTCAATGATGCAATTTGAAT 3900  
 QY 3850 GTGTGTAACAAGCTTTTCCGGCCCGGGTGAAGAGTGCATGCTTTATGGAATTAACCTA 3909  
 Db 3850 GTGTGTAACAAGCTTTTCCGGCCCGGGTGAAGAGTGCATGCTTTATGGAATTAACCTA 3909  
 QY 3901 GTGTGTAACAAGCTTTTCCGGCCCGGGTGAAGAGTGCATGCTTTATGGAATTAACCTA 3960  
 Db 3901 GTGTGTAACAAGCTTTTCCGGCCCGGGTGAAGAGTGCATGCTTTATGGAATTAACCTA 3960  
 QY 3910 TTAAGTGGCTTTATGAGTGCCTTCAAGAGCAATTTCTGAGCGAGACAGTGAAGAAATG 3969  
 Db 3910 TTAAGTGGCTTTATGAGTGCCTTCAAGAGCAATTTCTGAGCGAGACAGTGAAGAAATG 3969  
 QY 3961 TTAAGTGGCTTTATGAGTGCCTTCAAGAGCAATTTCTGAGCGAGACAGTGAAGAAATG 4020  
 Db 3961 TTAAGTGGCTTTATGAGTGCCTTCAAGAGCAATTTCTGAGCGAGACAGTGAAGAAATG 4020  
 QY 3970 GAATATGATCATGGAAGAACTCAACCTCTTATTTAGATTTCTGTAAATGAATCCAGTTTGT 4029  
 Db 3970 GAATATGATCATGGAAGAACTCAACCTCTTATTTAGATTTCTGTAAATGAATCCAGTTTGT 4029  
 QY 4021 GAATATGATCATGGAAGAACTCAACCTCTTATTTAGATTTCTGTAAATGAATCCAGTTTGT 4080  
 Db 4021 GAATATGATCATGGAAGAACTCAACCTCTTATTTAGATTTCTGTAAATGAATCCAGTTTGT 4080  
 QY 4030 GTTAATATTCCTGCTGAGAAACTTTAAAGTTTCAATATCCAAATGCTTCAAGTGTCC 4089  
 Db 4030 GTTAATATTCCTGCTGAGAAACTTTAAAGTTTCAATATCCAAATGCTTCAAGTGTCC 4089  
 QY 4081 GTTAATATTCCTGCTGAGAAACTTTAAAGTTTCAATATCCAAATGCTTCAAGTGTCC 4140  
 Db 4081 GTTAATATTCCTGCTGAGAAACTTTAAAGTTTCAATATCCAAATGCTTCAAGTGTCC 4140  
 QY 4090 ATAGTGAATTAATAGAGCCAGAGAGACAGCGGTGAGAGTTCAACTCTGAAGTGAAG 4149  
 Db 4090 ATAGTGAATTAATAGAGCCAGAGAGACAGCGGTGAGAGTTCAACTCTGAAGTGAAG 4149  
 QY 4141 ATAGTGAATTAATAGAGCCAGAGAGACAGCGGTGAGAGTTCAACTCTGAAGTGAAG 4200  
 Db 4141 ATAGTGAATTAATAGAGCCAGAGAGACAGCGGTGAGAGTTCAACTCTGAAGTGAAG 4200  
 QY 4150 CTGTCTCTCTGTGACCTTTGCAATGATGTCAGAGGCTGCGCTTGTGAGATGAGAGCT 4209  
 Db 4150 CTGTCTCTCTGTGACCTTTGCAATGATGTCAGAGGCTGCGCTTGTGAGATGAGAGCT 4209  
 QY 4201 CTGTCTCTCTGTGACCTTTGCAATGATGTCAGAGGCTGCGCTTGTGAGATGAGAGCT 4260  
 Db 4201 CTGTCTCTCTGTGACCTTTGCAATGATGTCAGAGGCTGCGCTTGTGAGATGAGAGCT 4260  
 QY 4210 TCAGGCTCCCATGATACATCTTCAGATCTTTTATTTTCAAGCTTCTGTACTTT 4269  
 Db 4210 TCAGGCTCCCATGATACATCTTCAGATCTTTTATTTTCAAGCTTCTGTACTTT 4269  
 QY 4261 TCAGGCTCCCATGATACATCTTCAGATCTTTTATTTTCAAGCTTCTGTACTTT 4320  
 Db 4261 TCAGGCTCCCATGATACATCTTCAGATCTTTTATTTTCAAGCTTCTGTACTTT 4320  
 QY 4270 ATGGAATTTGGCTG 4284  
 Db 4270 ATGGAATTTGGCTG 4284  
 QY 4321 ATGGAATTTGGCTG 4335  
 Db 4321 ATGGAATTTGGCTG 4335

## RESULT 12

AAD49436

ID AAD49436 standard; DNA; 4146 BP.

AC AAD49436;

DT 24-MAR-2003 (first entry)

DE Human blood cell surface antigen, CD109 encoding DNA #1.

KW Human; GPI-anchored TGF-beta1 binding protein; r150 protein; cancer;

KW glycosylphosphatidylinositol; transforming growth factor-beta1;

KW therapy; blood cell surface antigen; CD109; gene; ds.

OS Homo sapiens.

FH Key Location/Qualifiers

FT 1..4146

FT /tag= a

FT /product= "Human CD109 protein"





Db 1501 GAAAAATCTTGGACTCCAAAACCTGTGTAAATGTGATTAATATGAGATGATGGGAA 1560  
Qy 1524 ATTAATAGTATGTTCTTAAAAATTCCTGTCAGCTGTTTAAAAAATAAGATAAGCTA 1683  
Db 1561 ATTATTAAGTATGTTCTTAAAAATTCCTGTCAGCTGTTTAAAAAATAAGATAAGCTA 1620  
Qy 1684 TATTGAGTAAAGTAAAGCTGAAACCATCTGAGAAAGCTCTCTTAAAGATCTGTGACA 1743  
Db 1621 TATTGAGTAAAGTAAAGCTGAAACCATCTGAGAAAGCTCTCTTAAAGATCTGTGACA 1680  
Qy 1744 CAGGCTATCCGATAGTGGATGTGTAGCTGTGACAAAGTGGATCTGATGAATGCC 1803  
Db 1681 CAGGCTATCCGATAGTGGATGTGTAGCTGTGACAAAGTGGATCTGATGAATGCC 1740  
Qy 1804 TCTATGATATTACATGGAATAATGTGGTCCATGAGTTGAACTTTAATACAGATAT 1863  
Db 1741 TCTATGATATTACATGGAATAATGTGGTCCATGAGTTGAACTTTAATACAGATAT 1800  
Qy 1864 TATTAGGATGTTCAATGAAATCTTTTTCAGTCTTTTCAGAAATGTGAATCTGGGTAATG 1923  
Db 1801 TATTAGGATGTTCAATGAAATCTTTTTCAGTCTTTTCAGAAATGTGAATCTGGGTAATG 1860  
Qy 1924 ACAATGCAAACTTCAGAGAGATTTATTTGATGTGTTTATGCAATGCAATATGCT 1983  
Db 1861 ACAGATGCAAACTTCAGAGAGATTTATTTGATGTGTTTATGCAATGCAATATGCT 1920  
Qy 1984 GAGAGTTTATGAGAAATGGAAGACATATTTGATATTTCACTGACTTTCTTGGGT 2043  
Db 1921 GAGAGTTTATGAGAAATGGAAGACATATTTGATATTTCACTGACTTTCTTGGGT 1980  
Qy 2044 AGCAGTCCACATGTCGAAACATTTTTCAGAGACTTGGATTTGGCTAGACACCAATG 2103  
Db 1981 AGCAGTCCACATGTCGAAACATTTTTCAGAGACTTGGATTTGGCTAGACACCAATG 2040  
Qy 2104 GGTATCAGAGATTTACCAAGATTTGAAGTACTGATCTGATTTATCACTTCTGGGTG 2163  
Db 2041 GGTATCAGAGATTTACCAAGATTTGAAGTACTGATCTGATTTATCACTTCTGGGTG 2100  
Qy 2164 GCTACTGGTTTGTGATCTCTGAGGACTGTGGCTGTGACTTAACAACTACTCCAGTGGAG 2223  
Db 2101 GCTACTGGTTTGTGATCTCTGAGGACTGTGGCTGTGACTTAACAACTACTCCAGTGGAG 2160  
Qy 2224 CTCGAAGCTTCCAAACATTTTTCATTTTGTGAATCTTCCCTACTCTGTTATCAGAGT 2283  
Db 2161 CTCGAAGCTTCCAAACATTTTTCATTTTGTGAATCTTCCCTACTCTGTTATCAGAGT 2220  
Qy 2284 GAAAGATTTGCTTTGGAAATTAATCAATTAATTTGAAGATGCCATGAGTTAAG 2343  
Db 2221 GAAAGATTTGCTTTGGAAATTAATCAATTAATTTGAAGATGCCATGAGTTAAG 2280  
Qy 2344 GTAATCATGGAAGTGAACAAATTTGATTAATTTCAATGACTTCAAGTGAATTAATGCC 2403  
Db 2281 GTAATCATGGAAGTGAACAAATTTGATTAATTTCAATGACTTCAAGTGAATTAATGCC 2340  
Qy 2404 ACAGGCCACAGAGACCTTCTGTGTTCCAGTGAAGATGGGCAACTGTTCTTTTCCC 2463  
Db 2341 ACAGGCCACAGAGACCTTCTGTGTTCCAGTGAAGATGGGCAACTGTTCTTTTCCC 2400  
Qy 2464 ATCAGGCCACAACATCTGGAGAAATTCATCAAGTCAAGCTCTTTCACCACTGCT 2523  
Db 2401 ATCAGGCCACAACATCTGGAGAAATTCATCAAGTCAAGCTCTTTCACCACTGCT 2460  
Qy 2524 TCTGATGTGTACCCAGATGATTTTATGTAAGCTGAAGATGAAGAAATCAATTTCA 2583  
Db 2461 TCTGATGTGTACCCAGATGATTTTATGTAAGCTGAAGATGAAGAAATCAATTTCA 2520  
Qy 2584 CAATCCATCTTATTAAGACTGACAAATAGGCTACAGATACCTGAAAACCTTGGAGT 2643  
Db 2521 CAATCCATCTTATTAAGACTGACAAATAGGCTACAGATACCTGAAAACCTTGGAGT 2580  
Qy 2644 TTTCATATTTCTCTTAATACAGTGACTGGCAGTGAAGAGTTCAAGATCACTGCAATTGGA 2703  
Db 2581 TTTCATATTTCTCTTAATACAGTGACTGGCAGTGAAGAGTTCAAGATCACTGCAATTGGA 2640

Qy 2704 GATGTTCTTGCTGCTTCATCAATGAGCTTAGCCTCATGATTCGAGATGCTTAGT 2763  
Db 2641 GATGTTCTTGCTGCTTCATCAATGAGCTTAGCCTCATGATTCGAGATGCTTAGT 2700  
Qy 2764 GGTGAACAGAACATGATTAATTTTGTCTCCAAATATTTTACATTTTGGATTTATGACTTAA 2823  
Db 2701 GGTGAACAGAACATGATTAATTTTGTCTCCAAATATTTTACATTTTGGATTTATGACTTAA 2760  
Qy 2824 AAGAAACAATGACAGTATATTTGAAGAAAAGCTCTTCAATTAATGAGGCAAGGTTAC 2883  
Db 2761 AAGAAACAATGACAGTATATTTGAAGAAAAGCTCTTCAATTAATGAGGCAAGGTTAC 2820  
Qy 2884 CAGAGAACTTCTCATACAGAGAGAGATGCTCTTTCAGTCTTTTGGAAATTAAG 2943  
Db 2821 CAGAGAACTTCTCATACAGAGAGAGATGCTCTTTCAGTCTTTTGGAAATTAAG 2880  
Qy 2944 CTTCTGGAGACACTGTTGTGACCTTTTGTTTAAGATTTTCTTGAAGCCGATCCT 3003  
Db 2881 CTTCTGGAGACACTGTTGTGACCTTTTGTTTAAGATTTTCTTGAAGCCGATCCT 2940  
Qy 3004 TACATGATTTGATGAGAAATGTTTACAGAAACATACCTTGGCTTAAAGACATCAG 3063  
Db 2941 TACATGATTTGATGAGAAATGTTTACAGAAACATACCTTGGCTTAAAGACATCAG 3000  
Qy 3064 AAATCCAAACGGTGAATTTTGGATCCAGAGAGATTCATATGATGAGCTTCAAGGTGC 3123  
Db 3001 AAATCCAAACGGTGAATTTTGGATCCAGAGAGATTCATATGATGAGCTTCAAGGTGC 3060  
Qy 3124 AATTAAGTCCAGTAAACCTTACAGCTTATTTGTAATCTTCTCTGGGATTTAAG 3183  
Db 3061 AATTAAGTCCAGTAAACCTTACAGCTTATTTGTAATCTTCTCTGGGATTTAAG 3120  
Qy 3184 TATCAGCTTACATGATGTGCAAGAGTCAATCTTTTGGAGTGAATCAGTGA 3243  
Db 3121 TATCAGCTTACATGATGTGCAAGAGTCAATCTTTTGGAGTGAATCAGTGA 3180  
Qy 3244 GGAATTTACAGACATTAATCTAGGCTTATTAATCTTAATGATTCATCAGTGGAGT 3303  
Db 3181 GGAATTTACAGACATTAATCTAGGCTTATTAATCTTAATGATTCATCAGTGGAGT 3240  
Qy 3304 CTTAAAGCAGAGAGCTTTGAATATGCTGACTTGGAGACAGAACAGAGTGGCATG 3363  
Db 3241 CTTAAAGCAGAGAGCTTTGAATATGCTGACTTGGAGACAGAACAGAGTGGCATG 3300  
Qy 3364 CAATTCGGGTGTCAATCAGAGTCCAACTTTGCACTCCGGAGCAGCGCTCCCTGAT 3423  
Db 3301 CAATTCGGGTGTCAATCAGAGTCCAACTTTGCACTCCGGAGCAGCGCTCCCTGAT 3360  
Qy 3424 ATTGAAGTTCAGACCTATGACAGTCTCTCACATTTCAATTTTGAAGCTTGAAGGA 3483  
Db 3361 ATTGAAGTTCAGACCTATGACAGTCTCTCACATTTCAATTTTGAAGCTTGAAGGA 3420  
Qy 3484 ATCCCAATTAAGTGTGCTAAGCTAGCAGGCAAGAAATGCTTGGGTGTTTGCATCTA 3543  
Db 3421 ATCCCAATTAAGTGTGCTAAGCTAGCAGGCAAGAAATGCTTGGGTGTTTGCATCTA 3480  
Qy 3544 CAGGATACCACTGTGCTTTAAGGCTCTGTCTGAATTTGCACCCCTAATGAATACAGA 3603  
Db 3481 CAGGATACCACTGTGCTTTAAGGCTCTGTCTGAATTTGCACCCCTAATGAATACAGA 3540  
Qy 3604 AGGACAAATATCCAAAGTGAACCGTGAACGGGGCTTAGCTACCAAGTCTCTTGGTGTGA 3663  
Db 3541 AGGACAAATATCCAAAGTGAACCGTGAACGGGGCTTAGCTACCAAGTCTCTTGGTGTGA 3600  
Qy 3664 CAGCCAAATGCAATTAATTTCCGAATATGTTTGAATTTGATTTGTCAGCTCAAT 3723  
Db 3601 CAGCCAAATGCAATTAATTTCCGAATATGTTTGAATTTGATTTGTCAGCTCAAT 3660  
Qy 3724 GTTGTATTAATGTGAAGGCTTCTGGGTCTTCTGAAGAGAGATTTATCCAAAATCAA 3783  
Db 3661 GTTGTATTAATGTGAAGGCTTCTGGGTCTTCTGAAGAGAGATTTATCCAAAATCAA 3720



QY	3784	GAACCTTTGATTAGATGTTGCTGTAAAGAAAATAAAGATGATCTCAATCATGTGGAT	3843
Db	3721	GAACCTTTGATTAGATGTTGCTGTAAAGAAAATAAAGATGATCTCAATCATGTGGAT	3780
QY	3844	TTGAATGTGTACAAAGCTTTTTCGGGCCCGGGTGTAGAGTGGCATGGCTCTTATGAAAGTT	3903
Db	3781	TTGAATGTGTACAAAGCTTTTTCGGGCCCGGGTGTAGAGTGGCATGGCTCTTATGAAAGTT	3840
QY	3904	AACTATAATGAGCTTTATGATGCGCTTACAAGCAATTTCTGAGCGAACAAGTAAAG	3963
Db	3841	AACTATAATGAGCTTTATGATGCGCTTACAAGCAATTTCTGAGCGAACAAGTAAAG	3900
QY	3964	AAAGTGAATATGATTCATGAGAAAACCTCAACTATTATGATTCGTGAAATGAACCAG	4023
Db	3901	AAAGTGAATATGATTCATGAGAAAACCTCAACTATTATGATTCGTGAAATGAACCAG	3960
QY	4024	TTTTGTGTTAAATATCTCGCTGTGAGAACTTTAAAGTTTCAATAATCCCAAGATGCTTCA	4083
Db	3961	TTTTGTGTTAAATATCTCGCTGTGAGAACTTTAAAGTTTCAATAATCCCAAGATGCTTCA	4020
QY	4084	GTGTCATATGAGATTACTATGAGCCAAAGAGACAGCGGCGTGAAGAAAGTTACAATCTGAA	4143
Db	4021	GTGTCATATGAGATTACTATGAGCCAAAGAGACAGCGGCGTGAAGAAAGTTACAATCTGAA	4080
QY	4144	GTGAAGCTGTCTCTCTGTGACCTTTGCAAGTATGTCCAGGGCTCGCGTCTTGTGAGAGT	4203
Db	4081	GTGAAGCTGTCTCTCTGTGACCTTTGCAAGTATGTCCAGGGCTCGCGTCTTGTGAGAGT	4140
QY	4204	GGAGCT 4209	
Db	4141	GGAGCT 4146	

ID	AAID49437	standard; DNA; 4197 BP.
XX	AAID49437;	
XX	24-MAR-2003	(first entry)
DT	Human r150	DNA #2.
DE	Human; GPI-anchored TGF-beta1 binding protein; r150 protein; cancer	
KW	glycoeoylphosphatidyl inositol; transforming growth factor-beta1;	
KM	therapy; gene; ds.	
XX	Homo sapiens.	
OS		
Key	Location/Qualifiers	
FH	1..4197	
FT	/*tag= a	
FT	/product= "Human r1520 protein #3"	
FT	/transl_except= (pos:2044..2046, aa:Xaa)	
FT	/note= "Xaa corresponds to Ser, Tyr; No start	
FT	and stop codon"	
FT	/partial	
PN	WO200285942-A2.	
XX		
PD	31-OCT-2002.	
XX		
XX	24-APR-2002; 2002WO-CA00560.	
XX		
PR	24-APR-2001; 2001US-285713P.	
PR	14-FEB-2002; 2002US-356163P.	
XX		
PA	(UYMC-) UNIV MCGILL.	
XX		
PI	Phillip A, Tam B;	
XX		
XX	WPI, 2003-093100/08.	
DR	P-PSDB; AAE32015.	
DR		

XX Novel transforming growth factor (TGF)-beta 1 binding reagent which  
PT comprises r150 protein which acts as accessory receptor of TGF-beta,  
PT useful for negatively modulating TGF-beta activity, and thus for  
PT treating cancer -  
XX  
PS  
PS Claim 7; Page 118-120; 127pb; English.  
XX  
XX The invention relates to novel transforming growth factor (TGF)-beta1  
CC binding reagent which comprises a glycosylphosphatidyl inositol (GPI)-  
CC anchored TGF-beta1 binding protein referred to as r150 which acts as  
CC an accessory receptor of TGF-beta. The invention is used for negatively  
CC modulating TGF-beta activity, and thus for treating conditions  
CC characterised by overproduction of TGF-beta, such as cancer. Antisense  
CC molecules of the invention are used for increasing TGF-beta availability  
CC and increase graft success. The present sequence is human r150 protein  
CC encoding DNA.  
XX  
XX  
SQ Sequence 4197 BP; 1276 A; 788 C; 894 G; 1238 T; 1 other;

Query Match	95.2%;	Score 4082.4;	DB 25;	Length 4197;
Best Local Similarity	98.7%;	Pred. No. 0;		
Matches 4144;	Conservative	0;	Mismatches	2;
			Indels	51;
			Gaps	1;
QY	64	GTGCTCCCGGAGCTTCGGTTCCTGGTACAGAGCCCGAGATCATCAGGCGCGAGAAAT	123	
DB	1	GTGCTCCCGGAGCTTCGGTTCCTGGTACAGAGCCCGAGATCATCAGGCGCGAGAAAT	60	
QY	124	GTGACTATTTGGGGGTGAGCTTCTGGAACTCTGCCCTTACAGTGTCTGTGAAGGCGAG	183	
DB	61	GTGACTATTTGGGGGTGAGCTTCTGGAACTCTGCCCTTACAGTGTCTGTGAAGGCGAG	120	
QY	184	CTGCTCAAGACAGCATCAAACTCTGCTCTGTCTCTGTGAAGCAAGAGAGCTTTGAA	243	
DB	121	CTGCTCAAGACAGCATCAAACTCTGCTCTGTCTCTGTGAAGCAAGAGAGCTTTGAA	180	
QY	244	AAAGGCTCTTTAAGACACTTACTCTTCCATCACTACTCTGTGAACAGTGCAGATGAGATT	303	
DB	181	AAAGGCTCTTTAAGACACTTACTCTTCCATCACTACTCTGTGAACAGTGCAGATGAGATT	240	
QY	304	TATAGAGCTAGGTGAACCGGACGTAACCGAGATGAGATTTATTTCTTAATAGTACC	363	
DB	241	TATAGAGCTAGGTGAACCGGACGTAACCGAGATGAGATTTATTTCTTAATAGTACC	300	
QY	364	TTATCATTTGAGACCAAGAGAAATATCTGTCTCATTTAAACAGACAGAGCCTTATACAG	423	
DB	301	TTATCATTTGAGACCAAGAGAAATATCTGTCTCATTTAAACAGACAGAGCCTTATACAG	360	
QY	424	CCAAGACAGAGTGAAGTTCGATTCGATTTACACTCTTCCAGATTTTAAAGCTTTACAA	483	
DB	361	CCAAGACAGAGTGAAGTTCGATTCGATTTACACTCTTCCAGATTTTAAAGCTTTACAA	420	
QY	484	ACCTCTTTAAACATTTCTCATTTAAGAGCCCAATCAAAATTTGATCCAAAGTGTGTCA	543	
DB	421	ACCTCTTTAAACATTTCTCATTTAAGAGCCCAATCAAAATTTGATCCAAAGTGTGTCA	480	
QY	544	CAACAAAGTATCTTGGAGTCATTTCCAAACTTTTGAGTATTTTCCCATCAATACTT	603	
DB	481	CAACAAAGTATCTTGGAGTCATTTCCAAACTTTTGAGTATTTTCCCATCAATACTT	540	
QY	604	GGTACCTGTCGTATTTCAAGTTCAGTGAATGACAGACATTTTCAATCAATTTCAAGTT	663	
DB	541	GGTACCTGTCGTATTTCAAGTTCAGTGAATGACAGACATTTTCAATCAATTTCAAGTT	600	
QY	664	TCAGAAATATGTATTAACCAAAATTTGAAGTGAAGTTCGAGACACATTAATTTGTTCTATG	723	
DB	601	TCAGAAATATGTATTAACCAAAATTTGAAGTGAAGTTCGAGACACATTAATTTGTTCTATG	660	
QY	724	AATTTCAAGCATTTTAAATGTGATCCATCAAGCGCAAGTATTCATATGGGAAGCCAGTGA	783	
DB	661	AATTTCAAGCATTTTAAATGTGATCCATCAAGCGCAAGTATTCATATGGGAAGCCAGTGA	720	
QY	784	GGAGACGTAAACGTTACATTTTACCTTTATCTTTTGGGGAAGAAAGAAAATATTTACA	843	

```
Db 721 GGAAGCGTAACGCTTACATTTTATCTTATCTTTGGGAGAAAGAAAATTTTACA 780
Qy 844 AAAACATTTAAGATAAATGATCGCAAACTTCTCTTTAATGATGAAGATGAAAAAT 903
Db 781 AAAACATTTAAGATAAATGATCGCAAACTTCTCTTTAATGATGAAGATGAAAAAT 840
Qy 904 GTAATGATCTTCAATGATGACTTCTGATACCTGATCTATCTTCCCTGAGACAGTA 963
Db 841 GTAAAGATCTTCAATGATGACTTCTGATACCTGATCTATCTTCCCTGAGACAGTA 900
Qy 964 GAAATTTTAAACCAAGTACAGAAATCAGTTACAGGATTTTCAAGAAATGTAAGACTAAT 1023
Db 901 GAAATTTTAAACCAAGTACAGAAATCAGTTACAGGATTTTCAAGAAATGTAAGACTAAT 960
Qy 1024 GTGTTCTTCAAGCAACATGATTAACATCATGATTTTGTGATTAATGATTAATGATTAAG 1083
Db 961 GTGTTCTTCAAGCAACATGATTAACATCATGATTTTGTGATTAATGATTAATGATTAAG 1020
Qy 1084 CCATCTCTCAACTTCAACGCACTGTGAGGTAATCTGCTGATGAGCAACCACTGACT 1143
Db 1021 CCATCTCTCAACTTCAACGCACTGTGAGGTAATCTGCTGATGAGCAACCACTGACT 1080
Qy 1144 CTTGAAGAAAGAAATATATGATCTATAACAGTACACAGAGAACTATATCTGAGTAC 1203
Db 1081 CTTGAAGAAAGAAATATATGATCTATAACAGTACACAGAGAACTATATCTGAGTAC 1140
Qy 1204 TGGAGCGGATCTTAACATGGAATAATGAAATGAGAGCTTCCAGAAATTAATTAATCT 1263
Db 1141 TGGAGCGGATCTTAACATGGAATAATGAAATGAGAGCTTCCAGAAATTAATTAATCT 1200
Qy 1264 GTCCCCCAAGTGAACCTTTTAAGATTGAATTTCCCAATCTGAGAGATTTCCAGTACGTA 1323
Db 1201 GTCCCCCAAGTGAACCTTTTAAGATTGAATTTCCCAATCTGAGAGATTTCCAGTACGTA 1260
Qy 1324 CAGTTGAAGGCTTATTTCTTGTGATGAATAAGTACATGAGTTCATAGTCTGTTTAAAG 1383
Db 1261 CAGTTGAAGGCTTATTTCTTGTGATGAATAAGTACATGAGTTCATAGTCTGTTTAAAG 1320
Qy 1384 TCTCTCTAGTAGACATACATCAATCAATAAACAAGATGAAATTAATTAAGTGGATCG 1443
Db 1321 TCTCTCTAGTAGACATACATCAATCAATAAACAAGATGAAATTAATTAAGTGGATCG 1380
Qy 1444 CCTTTGAGTTGAGTGTAGTGAACAACAAGATTGAAGAGATTGAAGTATATGATGTA 1503
Db 1381 CCTTTGAGTTGAGTGTAGTGAACAACAAGATTGAAGAGATTGAAGTATATGATGTA 1440
Qy 1504 TCCAGGGGACAGTTGCTGCTGTAGAGAAACAAATTTCAACATGTTCTTTTAAACCA 1563
Db 1441 TCCAGGGGACAGTTGCTGCTGTAGAGAAACAAATTTCAACATGTTCTTTTAAACCA 1500
Qy 1564 GAAATTTTGGACTCCAAAAGCTGTATTTGTGATTAATTAAGATGAGGGGAA 1623
Db 1501 GAAATTTTGGACTCCAAAAGCTGTATTTGTGATTAATTAAGATGAGGGGAA 1560
Qy 1624 ATTAATGATGATTTCTTAAATTTCTGTTGAGGCTGTTTAAATTAATTAAGTAAAGTA 1683
Db 1561 ATTAATGATGATTTCTTAAATTTCTGTTGAGGCTGTTTAAATTAATTAAGTAAAGTA 1620
Qy 1684 TATTGAGTAAAGTAAAGTGAACCATCTGAGAAAGTCTCTCTAGAGATCTCTGTGACA 1743
Db 1621 TATTGAGTAAAGTAAAGTGAACCATCTGAGAAAGTCTCTCTAGAGATCTCTGTGACA 1680
Qy 1744 CAGCTGACTCCATAGTTGGGATTTGAGCTTTGACAAAGTGTGAATCTGATGAATGCC 1803
Db 1681 CAGCTGACTCCATAGTTGGGATTTGAGCTTTGACAAAGTGTGAATCTGATGAATGCC 1740
Qy 1804 TCTPATGATTTTAAACAATGGAATATGCTGATGAGTTGAACCTTTTAAACAAGATAT 1863
Db 1741 TCTPATGATTTTAAACAATGGAATATGCTGATGAGTTGAACCTTTTAAACAAGATAT 1800
Qy 1864 TATTAGGACATGTTCAATGATTTCTTTGCACTCTTTCAGAAATGAGACTCTGGGATATG 1923
Db 1801 TATTAGGACATGTTCAATGATTTCTTTGCACTCTTTCAGAAATGAGACTCTGGGATATG 1860
Qy 1924 ACAGATGCAAACTCCAGAAAGATTAATGATGATGTTTATGATGACATGACATATATGCT 1983
Db 1861 ACAGATGCAAACTCCAGAAAGATTAATGATGATGTTTATGATGATGATGACATATATGCT 1920
Qy 1984 GAGAGGTTTATGAGAGAAATGGAAGACATATTTGATGATTAATGATGATGATGATGATG 2043
Db 1921 GAGAGGTTTATGAGAGAAATGGAAGACATATTTGATGATGATGATGATGATGATGATG 1980
Qy 2044 AGCAGTCCACATGCTCCGAAAGCATTTTCCAGACATCTGATGATTTGGCTAGACACCAATG 2103
Db 1981 AGCAGTCCACATGCTCCGAAAGCATTTTCCAGACATCTGATGATTTGGCTAGACACCAATG 2040
Qy 2104 GGTATCAGATTTACCAAGAAATTTGAAGTAACTGTAATCTGATCTATCACTTCTTGGGATG 2163
Db 2041 GGTATCAGATTTACCAAGAAATTTGAAGTAACTGTAATCTGATCTATCACTTCTTGGGATG 2100
Qy 2164 GCTACTGATTTTGTGATCTCTGAGGACCTGGGATCTTGGACTTAAACAATCTGCACTGAG 2223
Db 2101 GCTACTGATTTTGTGATCTCTGAGGACCTGGGATCTTGGACTTAAACAATCTGCACTGAG 2160
Qy 2224 CTGAAAGCTTCCACATTTTTCATTTTGTGAATCTTCCCTACCTCTGATACAGAGT 2283
Db 2161 CTGAAAGCTTCCACATTTTTCATTTTGTGAATCTTCCCTACCTCTGATACAGAGT 2220
Qy 2284 GAAAGATTTGCTTGGAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2343
Db 2221 GAAAGATTTGCTTGGAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2280
Qy 2344 GTAATCATTTGAGAAAGTGAACAATTTGATTAATTAATTAATTAATTAATTAATTAATTA 2403
Db 2281 GTAATCATTTGAGAAAGTGAACAATTTGATTAATTAATTAATTAATTAATTAATTAATTA 2340
Qy 2404 ACAGGCGCACAGACACCTTCTGTTCCAGTGAAGATGAGGCAACCTGTTCTTTTCCC 2463
Db 2341 ACAGGCGCACAGACACCTTCTGTTCCAGTGAAGATGAGGCAACCTGTTCTTTTCCC 2400
Qy 2464 ATCAGGCGAACACATCTGGGAGAAATTTCTATCAGAGTCAACAGTCTTTCACCACTGCT 2523
Db 2401 ATCAGGCGAACACATCTGGGAGAAATTTCTATCAGAGTCAACAGTCTTTCACCACTGCT 2460
Qy 2524 TCTGATGCTGTCAACCAAGATTTTAAAGTCTGAAGAAATGAAGAAATCATATTTCA 2583
Db 2461 TCTGATGCTGTCAACCAAGATTTTAAAGTCTGAAGAAATGAAGAAATCATATTTCA 2520
Qy 2584 CAATCAATCTTATTAAGCTTGAAGTCAATAGGCTACAGATGACCTGAAACCTTTGAGT 2643
Db 2521 CAATCAATCTTATTAAGCTTGAAGTCAATAGGCTACAGATGACCTGAAACCTTTGAGT 2580
Qy 2644 TTCTCAATTTCTCTAATTAACAGTACCTGGCAGTGAAGAGTTGAGATCACTGCAATTTGA 2703
Db 2581 TTCTCAATTTCTCTAATTAACAGTACCTGGCAGTGAAGAGTTGAGATCACTGCAATTTGA 2640
Qy 2704 GATGTTCTTGTGCTTCTCAATCAATGAGCTTGAAGCTCATTTGATGAGTATGAGCTGT 2763
Db 2641 GATGTTCTTGTGCTTCTCAATCAATGAGCTTGAAGCTCATTTGATGAGTATGAGCTGT 2700
Qy 2764 GGTGAACAGAACATGATTAATTTGCTCCAAATATTTTGAATTAATCTGACTGATTA 2823
Db 2701 GGTGAACAGAACATGATTAATTTGCTCCAAATATTTTGAATTAATCTGACTGATTA 2760
Qy 2824 AAGAAACAATGACATATTAATTTGAAGAAAGCTCTTCAATTAATTAAGAGGAGTTAC 2883
Db 2761 AAGAAACAATGACATATTAATTTGAAGAAAGCTCTTCAATTAATTAAGAGGAGTTAC 2820
Qy 2884 CAGAGAACTTCTCATGAGAGAGAGATGCTCTTCAAGTCTTGGGATTTATGAC 2943
Db 2821 CAGAGAACTTCTCATGAGAGAGAGATGCTCTTCAAGTCTTGGGATTTATGAC 2880
Qy 2944 CTTCTGAGGAGACCTTGTGTTGAGCTTTTGAAGATTTTCTTGAAGCGATCT 3003
Db 2881 CTTCTGAGGAGACCTTGTGTTGAGCTTTTGAAGATTTTCTTGAAGCGATCT 2940
```

QY 3004 TACATGATATGATCAGATGTTGTTACACAGAACATACCTTGCTTAAGACATCAG 3063  
 DB 2941 TACATGATATGATCAGATGTTGTTACACAGAACATACCTTGCTTAAGACATCAG 3000  
 QY 3064 AATATCCAGGATGATTTGGATCCAGAGAGATGATTCATATGATGCTTCAAGGTGC 3123  
 DB 3001 AATATCCAGGATGATTTGGATCCAGAGAGATGATTCATATGATGCTTCAAGGTGC 3060  
 QY 3124 AATTAAGGTCAGTACACTTACAGCTATATTTGATCTCTCTCTGGATTTGAAG 3183  
 DB 3061 AATTAAGGTCAGTACACTTACAGCTATATTTGATCTCTCTCTGGATTTGAAG 3120  
 QY 3184 TATCAGCCATGATGATGTCAGAGATTCATCAATTTTGGAGCTGAATTCAGTAGA 3243  
 DB 3121 TATCAGCCATGATGATGTCAGAGATTCATCAATTTTGGAGCTGAATTCAGTAGA 3180  
 QY 3244 GGAATTTCAAGCAATTAATCTTACGCTTATTAATTAATGATGATGATGAGGAGT 3303  
 DB 3181 GGAATTTCAAGCAATTAATCTTACGCTTATTAATTAATGATGATGATGAGGAGT 3240  
 QY 3304 CCTAAGGAGAGAGCTTGAATATGCTGATGAGAGCAGAACAGAAAGGTGCATG 3363  
 DB 3241 CCTAAGGAGAGAGCTTGAATATGCTGATGAGAGCAGAACAGAAAGGTGCATG 3300  
 QY 3364 CAATTCGGGTGTCATCAGATGCCAACTTTCTGACTCCTGGCAGCAGCTCCCTGGAT 3423  
 DB 3301 CAATTCGGGTGTCATCAGATGCCAACTTTCTGACTCCTGGCAGCAGCTCCCTGGAT 3360  
 QY 3424 ATTTGAAGTTCAGACCTGATGCTCTGACACTTCTTCAATTTGACACTTGAAGGA 3483  
 DB 3361 ATTTGAAGTTCAGACCTGATGCTCTGACACTTCTTCAATTTGACACTTGAAGGA 3420  
 QY 3484 ATTTGAAGTTCAGACCTGATGCTCTGACACTTCTTCAATTTGACACTTGAAGGA 3543  
 DB 3421 ATTTGAAGTTCAGACCTGATGCTCTGACACTTCTTCAATTTGACACTTGAAGGA 3480  
 QY 3544 CAGATATACCTGCTGCTTTAAAGCTCTGCTGAAATTTGACAGCTTATGATACAGA 3603  
 DB 3481 CAGATATACCTGCTGCTTTAAAGCTCTGCTGAAATTTGACAGCTTATGATACAGA 3540  
 QY 3604 AGGACAAATATCCAGATGACCGTGACGGGGCTTACCTCAACAAGTCTCT 3651  
 DB 3541 AGGACAAATATCCAGATGACCGTGACGGGGCTTACCTCAACAAGTCTCT 3600  
 QY 3652 -----CTTCTGTGTGTAAGCCATG 3672  
 DB 3601 ATTTGACACACACACCGCTTACTCTTCAAGACAGAGACCTTGTGTGTAAGCCATG 3660  
 QY 3673 GCAGTTAATATTTCCGCAAAATGTTTGGATTTGCTATTTGCTAGCTCAATGTTGATAT 3732  
 DB 3661 GCAGTTAATATTTCCGCAAAATGTTTGGATTTGCTATTTGCTAGCTCAATGTTGATAT 3720  
 QY 3733 AATGTGAAGGCTTGTGGGTCTTCTAGAGACGAGATCTATCCAAATCAAGACCTTT 3792  
 DB 3721 AATGTGAAGGCTTGTGGGTCTTCTAGAGACGAGATCTATCCAAATCAAGACCTTT 3780  
 QY 3793 GATTTGATGTTGCTGTATAAAGAAATTAAGATGATCTCAATCATGTTGATTTGATG 3852  
 DB 3781 GATTTGATGTTGCTGTATAAAGAAATTAAGATGATCTCAATCATGTTGATTTGATG 3840  
 QY 3853 TGTACAGCTTTTGGGGCCGGGTGAGAGTGCATGAGCTCTTATGGAAGTTAACTATTA 3912  
 DB 3841 TGTACAGCTTTTGGGGCCGGGTGAGAGTGCATGAGCTCTTATGGAAGTTAACTATTA 3900  
 QY 3913 AGTGGCTTTATGATGCTTCAAGAGCAATTTCTGAGGAGACAGTGAAGAAAGTGA 3972  
 DB 3901 AGTGGCTTTATGATGCTTCAAGAGCAATTTCTGAGGAGACAGTGAAGAAAGTGA 3960  
 QY 3973 TATGATCATGAAAACTCAACTCTATTTAGATTTCTGTAATGAAACCCAGTTTGTGT 4032  
 DB 3961 TATGATCATGAAAACTCAACTCTATTTAGATTTCTGTAATGAAACCCAGTTTGTGT 4020

QY 4033 AATATTCCTGCTGTGAGAACTTTAAAGTTTCAATATCCAAATGCTTCAGTGTCCATA 4092  
 DB 4021 AATATTCCTGCTGTGAGAACTTTAAAGTTTCAATATCCAAATGCTTCAGTGTCCATA 4080  
 QY 4093 GTGATTTCTATGAGCCAGAGACAGCGGCTGAGAACTTCAACTCTGAAGTGAAGCTG 4152  
 DB 4081 GTGATTTCTATGAGCCAGAGACAGCGGCTGAGAACTTCAACTCTGAAGTGAAGCTG 4140  
 QY 4153 TCTCTCTGACCTTTTGAAGTATGTCAGAGGCTGCCCTCTGTGAGAGATGAGACT 4209  
 DB 4141 TCTCTCTGACCTTTTGAAGTATGTCAGAGGCTGCCCTCTGTGAGAGATGAGACT 4197  
 RESULT 14  
 AB079968  
 ID AB079968 standard; cDNA, 3535 BP.  
 XX  
 AC AB079968;  
 XX  
 DT 23-DEC-2002 (first entry)  
 XX  
 DE Human CD109 K15 protein encoding cDNA.  
 XX  
 KW CD109; alpha2 macroglobulin; thioester; cerebroprotective; cardiant;  
 KW immunosuppressive; haemostatic; anticoagulant; thrombolytic; human;  
 KW cardiovascular; vasotropic; gene therapy; CD109 K15; gene; ss.  
 OS Homo sapiens.  
 XX  
 XX Key Location/Qualifiers  
 FH 113..4450  
 FT /\*tag = a  
 FT /product = "CD109 K15"  
 FT  
 EN WO200270696-A2.  
 XX  
 PD 12-SEP-2002.  
 XX  
 PP 07-MAR-2002; 2002MO-CA00292.  
 XX  
 PR 07-MAR-2001; 2001US-273814P.  
 XX  
 PA (SCHU/) SCHUH A.  
 PA (SUTH/) SUTHERLAND R D.  
 XX  
 PI Schuh A, Sutherland RD;  
 XX  
 DR WPI; 2002-713450/77.  
 DR P-PEDB; ABB82169.  
 XX  
 PT New CD109 nucleic acids and polypeptides, useful in gene therapy,  
 PT particularly for treating strokes, myocardial infarctions, thrombosis,  
 PT thrombocytopenia, autoimmune diseases, or organ or bone marrow  
 PT transplantation -  
 XX  
 PS Claim 1; Fig 4a; 156pp; English.  
 XX  
 CC The invention relates to isolated nucleic acid molecules encoding CD109  
 CC polypeptides. These nucleic acid molecules include the human cDNA  
 CC sequences comprising CD109 K1, CD109 K1-H7, CD109 K15 or their variants.  
 CC CD109 is a novel member of the alpha2 macroglobulin (alpha2M)/C3, C4, C5  
 CC family of thioester-containing proteins. The CD109 polypeptides can be  
 CC expressed by standard recombinant methodology. The CD109 nucleic acid,  
 CC CD109 polypeptide or its mimetic, or the anti-CD109 antibody is useful  
 CC for treating or detecting a disease or disorder, e.g. conditions  
 CC associated with endothelial activation, platelet activation, activation  
 CC of the coagulation or fibrinolytic systems, activation of T lymphocytes  
 CC and of the complement system, quantitative or qualitative abnormalities  
 CC of platelet function, increased or impaired platelet aggregation and  
 CC activation, increased or impaired activation of the coagulation and/or  
 CC fibrinolytic systems, or impaired or increased immune activation. These  
 CC are also useful for treating cardiovascular disorders, stroke, myocardial  
 CC infarction, thrombosis, embolism, peripheral vascular disease,

CC thrombocytopenia, thrombocythemia, autoimmune diseases, organ  
CC transplantation, or bone marrow transplantation. The CD109 nucleic acids  
CC or their antisense nucleotide sequence are useful in gene therapy for  
CC treating these conditions. The present sequence represents the human  
CC CD109 K15 cDNA sequence.

XX Sequence 3535 BP; 1108 A; 675 C; 731 G; 1021 T; 0 other;

Query Match 74.4%; Score 3187.4; DB 24; Length 3535;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 3188; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGAGAGGCCACCGGCTCTGACCGCGCCCACTCTCTGCTGTCGACCGCGCGCTG 60  
DB 113 ATGAGAGGCCACCGGCTCTGACCGCGCCCACTCTCTGCTGTCGACCGCGCGCTG 172  
QY 61 GCCCGGCTCCCGGCGCTCTGCTGTCGACCGCGCCCACTCTCTGCTGTCGACCGCGCGCTG 120  
DB 173 GCCCGGCTCCCGGCGCTCTGCTGTCGACCGCGCCCACTCTCTGCTGTCGACCGCGCGCTG 232  
QY 121 AATGTAATTTGGGGTGGAGCTTCTGGGACATGCTGCTTCAAGTGAAGTGAAGGCG 180  
DB 233 AATGTAATTTGGGGTGGAGCTTCTGGGACATGCTGCTTCAAGTGAAGTGAAGGCG 292  
QY 181 GAGCTGCTCAAGACATCAACCTCACTGCTCTGCTCTGGAAGCAAGAGTCTTT 240  
DB 293 GAGCTGCTCAAGACATCAACCTCACTGCTCTGCTCTGGAAGCAAGAGTCTTT 352  
QY 241 GAAAAAGGCTCTTTAAGACACTTACTCTTCACTCACTCACTCACTCACTCACTCACTCACT 300  
DB 353 GAAAAAGGCTCTTTAAGACACTTACTCTTCACTCACTCACTCACTCACTCACTCACTCACT 412  
QY 301 ATTATGAGTACGCTGTAAACCGGACGTACCCAGATGAGATTTATTTCTTAATAGTACC 360  
DB 413 ATTATGAGTACGCTGTAAACCGGACGTACCCAGATGAGATTTATTTCTTAATAGTACC 472  
QY 361 CGCTTATCATTTGAGACCAAGAGAAATCTGTCTTCAATCAACAGACAAAGCCTTATAC 420  
DB 473 CGCTTATCATTTGAGACCAAGAGAAATCTGTCTTCAATCAACAGACAAAGCCTTATAC 532  
QY 421 AAGCCAAAGCAAGAGAAATTTGCGATTTGTTACATCTTCTCAAGATTTAAACCTTAC 480  
DB 533 AAGCCAAAGCAAGAGAAATTTGCGATTTGTTACATCTTCTCAAGATTTAAACCTTAC 592  
QY 481 AAAACCTCTTAAACATTTCTATTAAAGGACCCCAATCAATTTGATCCAAAGTGTG 540  
DB 593 AAAACCTCTTAAACATTTCTATTAAAGGACCCCAATCAATTTGATCCAAAGTGTG 652  
QY 541 TCACAAACAAGTGAATCTTGAAGTCAATTCCTTCAAGTATCTTCCATCAATA 600  
DB 653 TCACAAACAAGTGAATCTTGAAGTCAATTCCTTCAAGTATCTTCCATCAATA 712  
QY 601 CTGCTGATCTGCTTATTAAGTTCAGTGAATGACCAAGATTTATCAATCAATTTCAAG 660  
DB 713 CTGCTGATCTGCTTATTAAGTTCAGTGAATGACCAAGATTTATCAATCAATTTCAAG 772  
QY 661 GTTTCAGAAATATGATTAACCAAAATTTGAAGTGAATTTGAGACCAATTTATTTGTTCT 720  
DB 773 GTTTCAGAAATATGATTAACCAAAATTTGAAGTGAATTTGAGACCAATTTATTTGTTCT 832  
QY 721 ATGAATTTCAAGCAATTTAATGTAACATCAACGCAAGATTAATTAAGGACCAAGT 780  
DB 833 ATGAATTTCAAGCAATTTAATGTAACATCAACGCAAGATTAATTAAGGACCAAGT 892  
QY 781 AAAGAGAGCTTAAGCTTACATTTTACCTTTTACCTTTTGGGAAAGAGAAATATTT 840  
DB 893 AAAGAGAGCTTAAGCTTACATTTTACCTTTTACCTTTTGGGAAAGAGAAATATTT 952  
QY 841 ACAAAAACATTTAAGATTAATGATTCGCAAACTTCTCTTTAATGATGAAGAGATGAA 900  
DB 953 ACAAAAACATTTAAGATTAATGATTCGCAAACTTCTCTTTAATGATGAAGAGATGAA 1012  
QY 901 AATGTAATGATTCCTTCAATGAGACTTTCGAAATACCTGATCTATCTTCCCTGAGCA 960

DB 1013 AATGTAATGATTCCTTCAATGAGACTTTCGAAATACCTGATCTATCTTCCCTGAGCA 1072  
QY 961 GTGAAATTTTAAACACAGTACAGATCAGTATCAGTATTTCAAGAAATGTAAGACT 1020  
DB 1073 GTGAAATTTTAAACACAGTACAGATCAGTATCAGTATTTCAAGAAATGTAAGACT 1132  
QY 1021 AATGTTCTTCAAGCAACATGATTAATCAATTTGAGTATTTTATTAATTAATTAATTAAT 1080  
DB 1133 AATGTTCTTCAAGCAACATGATTAATCAATTTGAGTATTTTATTAATTAATTAATTAAT 1192  
QY 1081 AAGCCATCTTCAATTTCAAGCAACATGATTAATCAATTTGAGTATTTTATTAATTAATTAAT 1140  
DB 1193 AAGCCATCTTCAATTTCAAGCAACATGATTAATCAATTTGAGTATTTTATTAATTAATTAAT 1252  
QY 1141 ACTCTGAAAGAAAGAAATTAATGATTAATCAATTTGAGTATTTTATTAATTAATTAATTAAT 1200  
DB 1253 ACTCTGAAAGAAAGAAATTAATGATTAATCAATTTGAGTATTTTATTAATTAATTAATTAAT 1312  
QY 1201 TACTGAGCGGATCTAAACAGTGAATTAAGAAATGGAAGCTGTTCAAGAAATTAATTAAT 1260  
DB 1313 TACTGAGCGGATCTAAACAGTGAATTAAGAAATGGAAGCTGTTCAAGAAATTAATTAAT 1372  
QY 1261 ACTGTCCTTCAAGCAACATGATTAATCAATTTGAGTATTTTATTAATTAATTAATTAAT 1320  
DB 1373 ACTGTCCTTCAAGCAACATGATTAATCAATTTGAGTATTTTATTAATTAATTAATTAAT 1432  
QY 1321 CTACAGTGAAGGCTATTTCTTGGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAG 1380  
DB 1433 CTACAGTGAAGGCTATTTCTTGGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAG 1492  
QY 1381 AAGTCTCTAATGTAAGACATCAATCAATTAAGAAAGTAAAGTAAAGTAAAGTAAAGTAAAG 1440  
DB 1493 AAGTCTCTAATGTAAGACATCAATCAATTAAGAAAGTAAAGTAAAGTAAAGTAAAGTAAAG 1552  
QY 1441 TCGCTTTTGAATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTT 1500  
DB 1553 TCGCTTTTGAATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTT 1612  
QY 1501 GTATCCAGGAGGACAGTGTGAGTGTAGGAAACAAATTTCAACAAATTTCTCTTTTACA 1560  
DB 1613 GTATCCAGGAGGACAGTGTGAGTGTAGGAAACAAATTTCAACAAATTTCTCTTTTACA 1672  
QY 1561 CCAGAAATTTCTTGAATCTCCAAAGGCTGTATTTGTATTAATTAAGATGATGG 1620  
DB 1673 CCAGAAATTTCTTGAATCTCCAAAGGCTGTATTTGTATTAATTAAGATGATGG 1732  
QY 1621 GAAATTTAAGTATGTTCTTAAATTTCTGTTCACTTTGTTTAAATTAAGATTAAG 1680  
DB 1733 GAAATTTAAGTATGTTCTTAAATTTCTGTTCACTTTGTTTAAATTAAGATTAAG 1792  
QY 1681 CTATATTTGAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAG 1740  
DB 1793 CTATATTTGAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAG 1852  
QY 1741 ACAAGGCTGACCTCAATAGTGGATTTGATCTGTTGACAAAGATGATGATGAT 1800  
DB 1853 ACAAGGCTGACCTCAATAGTGGATTTGATCTGTTGACAAAGATGATGATGAT 1912  
QY 1801 GCTCTTAATGATTAATCAATGAAATTTGCTCATGATGATGATGATGATGATGATGATGATGAT 1860  
DB 1913 GCTCTTAATGATTAATCAATGAAATTTGCTCATGATGATGATGATGATGATGATGATGATGAT 1972  
QY 1861 TATTAATTTAGGAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAG 1920  
DB 1973 TATTAATTTAGGAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAG 2032  
QY 1921 TTGACAGATGCAAACTCAAGAGATTAATGATGATGATGATGATGATGATGATGATGATGAT 1980  
DB 2033 TTGACAGATGCAAACTCAAGAGATTAATGATGATGATGATGATGATGATGATGATGATGAT 2092  
QY 1981 GCTGAGAGTTTATGAGAGAAATGAGACATATTTGTATGATTTCTTTTCTTTG 2040

Db 2093 GCTGAGAGGTTTATGAGAGAAAATGAGACATATTGATGATTCATGACCTTTTCTTTG 2152  
 Qy 2041 GGTAGACGTCACATGTCGGAAGAGATTTTCCAGAGACTTGATTTGGCTAGACACAC 2100  
 Db 2153 GGTAGAGTCACATGTCGGAAGAGATTTTCCAGAGACTTGATTTGGCTAGACACAC 2212  
 Qy 2101 ATGGGTTACAGAGATTTACAGAGATTTGAGTACTGTAAGTATTCATCACTTTG 2160  
 Db 2213 ATGGGTTACAGAGATTTACAGAGATTTGAGTACTGTAAGTATTCATCACTTTG 2272  
 Qy 2161 GGGGCTACTGTTTGTGATCTCTGAGACCTGGGCTTGAGCTAACACTACTGAGT 2220  
 Db 2273 GGGGCTACTGTTTGTGATCTCTGAGACCTGGGCTTGAGCTAACACTACTGAGT 2332  
 Qy 2221 GAGCTCCAGACCTTCCAGACATTTTTCATTTTTCATCTCCCTACTCTGTTATCAGA 2280  
 Db 2333 GAGCTCCAGACCTTCCAGACATTTTTCATTTTTCATCTCCCTACTCTGTTATCAGA 2392  
 Qy 2281 GGTGAAGAAATTTGCTTGGAAAATTAATTAATTAATTAATTAATTAATTAATTAAT 2340  
 Db 2393 GGTGAAGAAATTTGCTTGGAAAATTAATTAATTAATTAATTAATTAATTAATTAAT 2452  
 Qy 2341 AAGGTATATATGAGAAAGTACCAATTTGATATCTTAATGACTTCAAGTAAATTAAT 2400  
 Db 2453 AAGGTATATATGAGAAAGTACCAATTTGATATCTTAATGACTTCAAGTAAATTAAT 2512  
 Qy 2401 GGCACAGGSCACAGAGACCTTCTGTTCCAGTGAAGAGGAGGCACTGTTCTTTT 2460  
 Db 2513 GGCACAGGSCACAGAGACCTTCTGTTCCAGTGAAGAGGAGGCACTGTTCTTTT 2572  
 Qy 2461 CCCATCAGGSCAACATCTGGAGAAATTCCTATCAGATCAGAGCTTTTCAACCACT 2520  
 Db 2573 CCCATCAGGSCAACATCTGGAGAAATTCCTATCAGATCAGAGCTTTTCAACCACT 2632  
 Qy 2521 GCTTCTGATGCTGTCAACCAAGATTTTATGTAAGGCTGGAAGAAATGAAATATCAT 2580  
 Db 2633 GCTTCTGATGCTGTCAACCAAGATTTTATGTAAGGCTGGAAGAAATGAAATATCAT 2692  
 Qy 2581 TCACATCATCTTATTAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 2640  
 Db 2693 TCACATCATCTTATTAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 2752  
 Qy 2641 AGTTTCTATTTCTCTCAATACAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 2700  
 Db 2753 AGTTTCTATTTCTCTCAATACAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 2812  
 Qy 2701 GGAAGATGTTCTGTTCTTCAATCAATGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 2760  
 Db 2813 GGAAGATGTTCTGTTCTTCAATCAATGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 2872  
 Qy 2761 TGGGTGAACAGAACATGATAATTTGCTCCAAATATTAATTTGATTAATCTGACT 2820  
 Db 2873 TGGGTGAACAGAACATGATAATTTGCTCCAAATATTAATTTGATTAATCTGACT 2932  
 Qy 2821 AAAAAGAAACAACTGACAGATTAATTTGAAGAAAAGCTTTTCAATTTTGAAGGCAAGT 2880  
 Db 2933 AAAAAGAAACAACTGACAGATTAATTTGAAGAAAAGCTTTTCAATTTTGAAGGCAAGT 2992  
 Qy 2881 TACCAAGAGAACTTCTATCAAGAGGAAAGTGGCTTTTCAAGTCTTTTGGAAATTAAT 2940  
 Db 2993 TACCAAGAGAACTTCTATCAAGAGGAAAGTGGCTTTTCAAGTCTTTTGGAAATTAAT 3052  
 Qy 2941 GACCCCTTGGGAGACCTTGGTTCAGCTTTTGTAAAGATGTTTCTTGAAGCCGAT 3000  
 Db 3053 GACCCCTTGGGAGACCTTGGTTCAGCTTTTGTAAAGATGTTTCTTGAAGCCGAT 3112  
 Qy 3001 CCTTACATGATTAATGATGAGATGTTTACACAGACATCACTTGGCTTAAAGACAT 3060  
 Db 3113 CCTTACATGATTAATGATGAGATGTTTACACAGACATCACTTGGCTTAAAGACAT 3172  
 Qy 3061 CAGAAATCCACCGTGAATTTTGGATCCAGAGAGATGATTCATAGTAGCTTCAAGT 3120  
 Db 3173 CAGAAATCCACCGTGAATTTTGGATCCAGAGAGATGATTCATAGTAGCTTCAAGT 3232

Qy 3121 GGCATTAAGATCCAGTAACCTTACAGCCTATATTTGATCTTCTCTGGATATAGA 3180  
 Db 3223 GGCATTAAGATCCAGTAACCTTACAGCCTATATTTGATCTTCTCTGGATATAGA 3292  
 Qy 3181 AAGTATCAG 3189  
 Db 3293 AAGTATCAG 3301  
 RESULT 15  
 ABQ79969  
 ID ABQ79969 standard; cDNA; 3535 BP.  
 XX  
 AC ABQ79969;  
 XX  
 DT 23-DEC-2002 (first entry)  
 XX  
 DE Human CD109 K15 variant protein encoding cDNA.  
 XX  
 KW CD109; alpha2 macroglobulin; thioester; cerebroprotective; cardiac;  
 KW immunosuppressive; haemostatic; anticoagulant; thrombolytic; human;  
 KW cardiovascular; vasotropic; gene therapy; CD109 K15; variant; gene; ss.  
 OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 FT CDS 113..4450  
 FT /tag= a  
 FT /product= "CD109 K15 variant"  
 PN W0200270696-A2.  
 PD 12-SEP-2002.  
 XX 07-MAR-2002; 2002WO-CA00292.  
 XX 07-MAR-2001; 2001US-273814P.  
 PR (SCHU/) SUTHERLAND R D.  
 PA (SUTR/) SUTHERLAND R D.  
 XX Schuh A, Sutherland RD;  
 PI MPI. 2002-713450/77.  
 DR P-ESDB; ABB82170.  
 XX  
 PT New CD109 nucleic acids and polypeptides, useful in gene therapy,  
 PT particularly for treating strokes, myocardial infarctions, thrombosis,  
 PT thrombocytopenia, autoimmune diseases, or organ or bone marrow  
 PT transplantation -  
 XX  
 XX Claim 1; Fig 4b; 156bp; English.  
 XX  
 CC The invention relates to isolated nucleic acid molecules encoding CD109  
 CC polypeptides. These nucleic acid molecules include the human cDNA  
 CC sequences comprising CD109 K1, CD109 K1-H7, CD109 K15 or their variants.  
 CC CD109 is a novel member of the alpha2 macroglobulin (alpha2M)/G3, C4, C5  
 CC family of thioester-containing proteins. The CD109 polypeptides can be  
 CC expressed by standard recombinant methodology. The CD109 nucleic acid,  
 CC CD109 polypeptide or its mimetic, or the anti-CD109 antibody is useful  
 CC for treating or detecting a disease or disorder, e.g. conditions  
 CC associated with endothelial activation, platelet activation, activation  
 CC of the coagulation or fibrinolytic systems, activation of T lymphocytes  
 CC and of the complement system, quantitative or qualitative abnormalities  
 CC of platelet function, increased or impaired platelet aggregation and  
 CC activation, increased or impaired activation of the coagulation and/or  
 CC fibrinolytic systems, or impaired or increased immune activation. These  
 CC are also useful for treating cardiovascular disorders, stroke, myocardial  
 CC infarction, thrombosis, embolism, peripheral vascular disease,  
 CC thrombocytopenia, thrombocytopenia, autoimmune diseases, organ  
 CC transplantation, or bone marrow transplantation. The CD109 nucleic acids  
 CC or their antisense nucleotide sequence are useful in gene therapy for

CC treating these conditions. The present sequence represents the human  
CD109 K15 variant cDNA sequence.

XX Sequence 3535 BP; 1107 A; 676 C; 731 G; 1021 T; 0 other;

Query Match 74.3%; Score 3185.8; DB 24; Length 3535;

Beet Local Similarity 99.9%; Fred. No. 0;

Matches 3187; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
OY 1 ATGAGAGGCGCCAGCTCTGACGCGCGCCAGCTCTGCGGTGTGACCGCGCGCTG 60
DB 113 ATGAGAGGCGCCAGCTCTGACGCGCGCCAGCTCTGCGGTGTGACCGCGCGCTG 172
OY 61 GCCGTGCTCCCGGCGCTCGGTTTCTGTGTGACAGCGCCAGGATCATCAGGCCGAGAGA 120
DB 173 GCCGTGCTCCCGGCGCTCGGTTTCTGTGTGACAGCGCCAGGATCATCAGGCCGAGAGA 232
OY 121 AATGTGACTATTGGGGGTGAGACTTGTGACACACTGCTTCAACAGGTACTGTGAGGCG 180
DB 223 AATGTGACTATTGGGGGTGAGACTTGTGACACACTGCTTCAACAGGTACTGTGAGGCG 292
OY 181 GAGCTGCTCAAGACAGCATCAAACTCACTGCTCTGCTGTGAGAGCAGAGAGGCTTT 240
DB 293 GAGCTGCTCAAGACAGCATCAAACTCACTGCTCTGCTGTGAGAGCAGAGAGGCTTT 352
OY 241 GAAAAAGGCTCTTTTAAGACACTTACTCTTCATCTACTACTCTGAAACAGTCAAGATAG 300
DB 353 GAAAAAGGCTCTTTTAAGACACTTACTCTTCATCTACTACTCTGAAACAGTCAAGATAG 412
OY 301 AATTATGAGCTACGTGTACCGGAGCTACCGAGATGAGATTTATTTCTTAATAGTACC 360
DB 413 AATTATGAGCTACGTGTACCGGAGCTACCGAGATGAGATTTATTTCTTAATAGTACC 472
OY 361 CGCTTATGCTTGAAGACCAAGAGAAATATCTGCTTCAATCAACAGCAAGGCTTATAC 420
DB 473 CGCTTATGCTTGAAGACCAAGAGAAATATCTGCTTCAATCAACAGCAAGGCTTATAC 532
OY 421 AAGCCAAAGCAAGAGAGTTCGCAATGTTCACCTCTCTGAGATTTTAAAGCTTAC 480
DB 533 AAGCCAAAGCAAGAGAGTTCGCAATGTTCACCTCTCTGAGATTTTAAAGCTTAC 592
OY 481 AAAACCTTTTAAACATTTCTCAATTAAGAACCCCAATTAATTTGATCAACAGTGTG 540
DB 593 AAAACCTTTTAAACATTTCTCAATTAAGAACCCCAATTAATTTGATCAACAGTGTG 652
OY 541 TCACAAAGAGTGAATCTGAGTCAATTCGAAACCTTTGAGTATCTCCATCCATA 600
DB 653 TCACAAAGAGTGAATCTGAGTCAATTCGAAACCTTTGAGTATCTCCATCCATA 712
OY 601 CTGTGTCAGTGTCTATCAAGTTCAGTGAATGACAGACATTAATCAATCAATTCAG 660
DB 713 CTGTGTCAGTGTCTATCAAGTTCAGTGAATGACAGACATTAATCAATCAATTCAG 772
OY 661 GTTTCAGAAATATGATTCGAAATTTGAAGTGAAGTTCGACAGACCACTTAATTTGTTCT 720
DB 773 GTTTCAGAAATATGATTCGAAATTTGAAGTGAAGTTCGACAGACCACTTAATTTGTTCT 832
OY 721 ATGATTTGAAGCAATTAATGATGATCAATCAAGCAAGTATATGAGGAGAGCGAGT 780
DB 833 ATGATTTGAAGCAATTAATGATGATCAATCAAGCAAGTATATGAGGAGAGCGAGT 892
OY 781 AAGAGAGAGCTTAAGCTTACATTTTATCTTATCTTTGGGAAAGAAATATTT 840
DB 893 AAGAGAGAGCTTAAGCTTACATTTTATCTTATCTTTGGGAAAGAAATATTT 952
OY 841 ACAAACACATTTAAGATTAATGATCTGCAAACTTCTCTTTTAATGATGAAGATGAAA 900
DB 953 ACAAACACATTTAAGATTAATGATCTGCAAACTTCTCTTTTAATGATGAAGATGAAA 1012
OY 901 AATGTAATGATTTCTTGAATGAGCTTTCTGAATATCTGATCTATCTTCCCTGAGACA 960
DB 1013 AATGTAATGATTTCTTGAATGAGCTTTCTGAATATCTGATCTATCTTCCCTGAGACA 1072
```

```
OY 961 GTAGAAATTTTAAACACAGTGAAGAAATCACTTACAGATTTTTCAGAAATGTAAGCACT 1020
DB 1073 GTAGAAATTTTAAACACAGTGAAGAAATCACTTACAGATTTTTCAGAAATGTAAGCACT 1132
OY 1021 AATGTGCTTCAAGCAACATGATTCATCATGAGTTTATGATTAATATCTACTGCTTG 1080
DB 1133 AATGTGCTTCAAGCAACATGATTCATCATGAGTTTATGATTAATATCTACTGCTTG 1192
OY 1081 AAGCATCTCTCAACTTCAAGCCACTGTGAAAGTAACTGTCGTGAGCAACCAACTG 1140
DB 1193 AAGCATCTCTCAACTTCAAGCCACTGTGAAAGTAACTGTCGTGAGCAACCAACTG 1252
OY 1141 ACTTGAAGAAAGAAATTAATGATGATTAACAGTGAACACAGAGAAATTAATCTGAG 1200
DB 1253 ACTTGAAGAAAGAAATTAATGATGATTAACAGTGAACACAGAGAAATTAATCTGAG 1312
OY 1201 TACTGAGCGGATCTTAACAGTGAAGTCAAGAAATGGAAGCTGTCAGAAATTAATAT 1260
DB 1313 TACTGAGCGGATCTTAACAGTGAAGTCAAGAAATGGAAGCTGTCAGAAATTAATAT 1372
OY 1261 ACTGCCCCCAAGTGAACCTTTTAAGATTTGAATTCCTCAATCTGAGAGATTCAGTGAG 1320
DB 1373 ACTGCCCCCAAGTGAACCTTTTAAGATTTGAATTCCTCAATCTGAGAGATTCAGTGAG 1432
OY 1321 CTACAGTTGAAGGCTTATTTCTTGTGATGAAGTGAAGTCAAGTCAATGCTGTTT 1380
DB 1433 CTACAGTTGAAGGCTTATTTCTTGTGATGAAGTGAAGTCAAGTCAATGCTGTTT 1492
OY 1381 AAGTCTCTGATGAACATCATCCAACTTAACAAAGAGATGAAGTAAAGTGGGA 1440
DB 1493 AAGTCTCTGATGAACATCATCCAACTTAACAAAGAGATGAAGTAAAGTGGGA 1552
OY 1441 TCGCCTTTTGTGTTGTGTTAGTGTGCAACAAAGATTTGAAGATTTAAGTATGTA 1500
DB 1553 TCGCCTTTTGTGTTGTGTTAGTGTGCAACAAAGATTTGAAGATTTAAGTATGTA 1612
OY 1501 GTATCCAGGGGACAGTTGTGTGTGAGAAACAAATTCACAAATGTTCTCTTAAACA 1560
DB 1613 GTATCCAGGGGACAGTTGTGTGTGAGAAACAAATTCACAAATGTTCTCTTAAACA 1672
OY 1561 CCAAGAAATCTTGTGACTCCAAAGCTGTGTAATTTGTATTAATTAAGATGATGG 1620
DB 1673 CCAAGAAATCTTGTGACTCCAAAGCTGTGTAATTTGTATTAATTAAGATGATGG 1732
OY 1621 GAAATTAATGATGATCTTAAATTTCCGTGTCACTGTTTAAATTAATGAATGAAG 1680
DB 1733 GAAATTAATGATGATCTTAAATTTCCGTGTCACTGTTTAAATTAATGAATGAAG 1792
OY 1681 CTATATTTGAAGTAAAGTGAAGCTGAACCATCTGAGAAAGTCTCTTGAAGATCTGTG 1740
DB 1793 CTATATTTGAAGTAAAGTGAAGCTGAACCATCTGAGAAAGTCTCTTGAAGATCTGTG 1852
OY 1741 ACAACGCTGATCTCCATAGTTGGGATTTGATGCTGTGACAAAGTGTGAATCTGATGAT 1800
DB 1853 ACAACGCTGATCTCCATAGTTGGGATTTGATGCTGTGACAAAGTGTGAATCTGATGAT 1912
OY 1801 GCCCTTAATGATTAATCAATGAAATGTGTCCATGATGTGGAACCTTATTAACAGAGA 1860
DB 1913 GCCCTTAATGATTAATCAATGAAATGTGTCCATGATGTGGAACCTTATTAACAGAGA 1972
OY 1861 TATTAATTTAGGAGTGTCAATGAATTTCTTTGAGTCTTCAAGAAATGAGACTGTGGTA 1920
DB 1973 TATTAATTTAGGAGTGTCAATGAATTTCTTTGAGTCTTCAAGAAATGAGACTGTGGTA 2032
OY 1921 TTGACAGATGCAAACTCAAGAGATTAATTTGATGATGTTTATTAACATGAGAAATAT 1980
DB 2033 TTGACAGATGCAAACTCAAGAGATTAATTTGATGATGTTTATTAACATGAGAAATAT 2092
OY 1981 GCTGAGAGGTTTATGAGAGAAATGGAAGCATATGATGATTAATGATGCTTTCTTTG 2040
DB 2093 GCTGAGAGGTTTATGAGAGAAATGGAAGCATATGATGATTAATGATGCTTTCTTTG 2152
OY 2041 GGTAGAGTCCACATGTCGAAAGCATTTTCCAGAGACTTGTGATTTGGCTGAGACCAAC 2100
```



Db 2153 GGTAGAGTCCATGTCGGAAGCATTTTCCAGACCTTGATTTGGCTAGACACAAAC 2212  
Qy 2101 ATGGGTTACAGATTTTACCAAGAAATTTGAAGTAACTGTACTGATTTCTATCACTTCTGG 2160  
Db 2213 ATGGGTTCCAGATTTTACCAAGAAATTTGAAGTAACTGTACTGATTTCTATCACTTCTGG 2272  
Qy 2161 GTGGCTACTGGTTTGTGATCTCTGAGAGACTGGGTTCTTGAATCAACTCTGATGAGT 2220  
Db 2273 GTGGCTACTGGTTTGTGATCTCTGAGAGACTGGGTTCTTGAATCAACTCTGATGAGT 2332  
Qy 2221 GAGCTTCAAGCTTCCAACTTTTTCATTTTGAATCTTCCCTACTCTGTTATCAGA 2280  
Db 2333 GAGCTTCAAGCTTCCAACTTTTTCATTTTGAATCTTCCCTACTCTGTTATCAGA 2392  
Qy 2281 GGTGAAGAAATTTGCTTGGAAATTAATTAATTAATTAATTAATTAATTAATTAAT 2340  
Db 2393 GGTGAAGAAATTTGCTTGGAAATTAATTAATTAATTAATTAATTAATTAATTAAT 2452  
Qy 2341 AAGGTAAATCATGAGAAAGTGACAAATTTGATATCTAATGACTTCAAGTAAATTAAT 2400  
Db 2453 AAGGTAAATCATGAGAAAGTGACAAATTTGATATCTAATGACTTCAAGTAAATTAAT 2512  
Qy 2401 GCCACAGGCCAACCCAGACCTTCTGTTCCAGTGAAGATGGGCAACTGTTCTTTT 2460  
Db 2513 GCCACAGGCCAACCCAGACCTTCTGTTCCAGTGAAGATGGGCAACTGTTCTTTT 2572  
Qy 2461 CCCATAGGCGCAACATCTGGAGAAATTCCTATGACAGTCAAGCTCTTCAACCACT 2520  
Db 2573 CCCATAGGCGCAACATCTGGAGAAATTCCTATGACAGTCAAGCTCTTCAACCACT 2632  
Qy 2521 GCTTCTGATGCTGTCAACCCAGATGATTTTGTAAAGGCTGAAGGAATAGAAAAATCATAT 2580  
Db 2633 GCTTCTGATGCTGTCAACCCAGATGATTTTGTAAAGGCTGAAGGAATAGAAAAATCATAT 2692  
Qy 2581 TCAATATCATCTTATTTAAGCTTGAATGATGATGATGATGATGATGATGATGATGAT 2640  
Db 2693 TCAATATCATCTTATTTAAGCTTGAATGATGATGATGATGATGATGATGATGATGAT 2752  
Qy 2641 AGTTTCTATTTCTCTCTAATCAGTGAATGAGGAGGAAAGGTTGAGATCACTGCAAT 2700  
Db 2753 AGTTTCTATTTCTCTCTAATCAGTGAATGAGGAGGAAAGGTTGAGATCACTGCAAT 2812  
Qy 2701 GAGATGTTCTTGTCTTCTCAATCAATGCTTGAAGCTTCAATGCTTGAAGCTTGAAGCT 2760  
Db 2813 GAGATGTTCTTGTCTTCTCAATCAATGCTTGAAGCTTCAATGCTTGAAGCTTGAAGCT 2872  
Qy 2761 TGTGTGAACAGAACTGATTAATTTTGTCTCAAAATATTTAATTTGATTAATCTGACT 2820  
Db 2873 TGTGTGAACAGAACTGATTAATTTTGTCTCAAAATATTTAATTTGATTAATCTGACT 2932  
Qy 2821 AAAAAAGAACTGACAGATTAATTTGAAGAAAGCTCTTCAATTTATGAGCAAGGT 2880  
Db 2933 AAAAAAGAACTGACAGATTAATTTGAAGAAAGCTCTTCAATTTATGAGCAAGGT 2992  
Qy 2881 TACACAGAGAACTTCTATCAGAGGAAAGATGCTCTTCAAGCTTGGGAAATTAAT 2940  
Db 2993 TACACAGAGAACTTCTATCAGAGGAAAGATGCTCTTCAAGCTTGGGAAATTAAT 3052  
Qy 2941 GACCTTCTGGAGCACTTGTGTGAGCTTGTGTTAAGATGTTCTTGAAGCCGAT 3000  
Db 3053 GACCTTCTGGAGCACTTGTGTGAGCTTGTGTTAAGATGTTCTTGAAGCCGAT 3112  
Qy 3001 CCTTACATGATTTGATGAGATGTTGTTACAGAACTACACTTGGCTTAAAGGACAT 3060  
Db 3113 CCTTACATGATTTGATGAGATGTTGTTACAGAACTACACTTGGCTTAAAGGACAT 3172  
Qy 3061 CAGAAATCAAGGTAATTTTGGATCCAGAAAGATGATTCATGATGAGCTTCAAGGT 3120  
Db 3173 CAGAAATCAAGGTAATTTTGGATCCAGAAAGATGATTCATGATGAGCTTCAAGGT 3232  
Qy 3121 GGCATATAAAGTCCAGTAACTTACAGCTATATTTGTAATTTCTCTCTGGATATAGA 3180

Db 3233 GGCATATAAAGTCCAGTAACTTACAGCTTATATTTGTAATTTCTCTCTGGATATAGA 3292  
Qy 3181 AAGTATCAG 3189  
Db 3293 AAGTATCAG 3301

Search completed: January 16, 2004, 14:38:44  
Job time : 1046 secs

**THIS PAGE BLANK (08970)**

1	4281.2	99.9	4369	25	AA049434
2	4224.4	98.5	4761	24	AB079964
3	4224.4	98.5	5882	25	AA049440
4	4224.4	99.5	5895	24	AB079961
5	4223.4	98.5	5883	25	AA049435
6	4222.8	98.5	4761	24	AB079965
7	4222.8	98.5	5895	24	AB079967
8	4221.4	98.5	4335	24	AA149815
					Human platelet all
					Human r150 gene #3
					Human CD109 K1 pre
					Human blood cell
					Human CD109 K1-H7
					Human r150 DNA #1
					Human CD109 K1 val
					Human CD109 K1-H7
					Human platelet all

PR 24-APR-2001; 2001US-285713P  
DE 14 FEB 2003 09:51:53Z

PR 14-FEB-2002; 2002US-356163P

**THIS PAGE BLANK (USPTO)**

XX (UWMC-) UNIV MCGILL.  
PA Philip A, Tam B;  
PI WPI, 2003-093100/08.  
XX P-PSDB; AAB32012.

DR Novel transforming growth factor (TGF)-beta 1 binding reagent which  
PT comprises r150 protein which acts as accessory receptor of TGF-beta,  
PT useful for negatively modulating TGF-beta activity, and thus for  
PT treating cancer

PS Claim 7, Page 91-93; 127p; English.

XX The invention relates to novel transforming growth factor (TGF)-beta 1  
XX binding reagent which comprises a glycosylphosphatidyl inositol (GPI)-  
XX anchored TGF-beta binding protein referred to as r150 which acts as  
XX an accessory receptor of TGF-beta. The invention is used for negatively  
XX modulating TGF-beta activity, and thus for treating conditions  
XX characterised by overproduction of TGF-beta, such as cancer. Antisense  
XX molecules of the invention are used for increasing TGF-beta availability  
XX and increase graft success. The present sequence is a gene encoding  
XX human r150 protein.

XX Sequence 4369 BP; 1301 A; 842 C; 932 G; 1293 T; 1 other;

Query Match 99.9%; Score 4281.2; DB 25; Length 4369;

Best Local Similarity 99.9%; Pred. No. 0;

Matches 4283; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 ATGAGAGGCGCCACCGCTCTGACCGCGCCACCTCTGCGGTGACCGCGCGCTG 60  
DB 26 ATGAGAGGCGCCACCGCTCTGACCGCGCCACCTCTCTGCGGTGACCGCGCGCTG 85  
QY 61 GCGGTGCTCCCGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 120  
DB 86 GCGGTGCTCCCGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 145  
QY 121 AATGTACTATTGGGGGTGAGCTTCTGAACTGCTGCTGCTGCTGCTGCTGCTGCTG 180  
DB 146 AATGTACTATTGGGGGTGAGCTTCTGAACTGCTGCTGCTGCTGCTGCTGCTGCTG 205  
QY 181 GAGTGTCTCAAGACAGCATCAAACTCACTGCTCTGCTGCTGCTGCTGCTGCTGCTG 240  
DB 206 GAGTGTCTCAAGACAGCATCAAACTCACTGCTCTGCTGCTGCTGCTGCTGCTGCTG 265  
QY 241 GAAAAAGCTCTTTTAAGACACTTACTCTTCCATCACTGCTGCTGCTGCTGCTGCTG 300  
DB 266 GAAAAAGCTCTTTTAAGACACTTACTCTTCCATCACTGCTGCTGCTGCTGCTGCTG 325  
QY 301 AATTATGAGCTACGCTGACCGGAGTACCGGAGTACCGGAGTACCGGAGTACCGGAGT 360  
DB 326 AATTATGAGCTACGCTGACCGGAGTACCGGAGTACCGGAGTACCGGAGTACCGGAGT 385  
QY 361 CGCTTATCAATTGAGACCAAGAGATATGCTTCTGCTGCTGCTGCTGCTGCTGCTGCT 420  
DB 386 CGCTTATCAATTGAGACCAAGAGATATGCTTCTGCTGCTGCTGCTGCTGCTGCTGCT 445  
QY 421 AAGCCAAAGCAAGAGTGAAGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 480  
DB 446 AAGCCAAAGCAAGAGTGAAGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 505  
QY 481 AAAAAGCTCTTTAAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 540  
DB 506 AAAAAGCTCTTTAAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 565  
QY 541 TCAACAACAAGTATCTGAGAGTATTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 600  
DB 566 TCAACAACAAGTATCTGAGAGTATTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 625  
QY 601 CTGAGTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 660

DB 626 CTGAGTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 685  
QY 661 GTTTCAGAAATATGATTAACCAAAATTTGAAGTACCTTGGAGACACCATTAATATGCT 720  
DB 686 GTTTCAGAAATATGATTAACCAAAATTTGAAGTACCTTGGAGACACCATTAATATGCT 745  
QY 721 ATGAATTTAAGCATTTAATGTAACCATGCGGCAAGTATATATGGAAGCCAGTG 780  
DB 746 ATGAATTTAAGCATTTAATGTAACCATGCGGCAAGTATATATGGAAGCCAGTG 805  
QY 781 AAAGAGACGTAAGCTTACCTTTTACCTTTTACCTTTTACCTTTTACCTTTTACCTTT 840  
DB 806 AAAGAGACGTAAGCTTACCTTTTACCTTTTACCTTTTACCTTTTACCTTTTACCTTT 865  
QY 841 ACAAATTAATTAAGTAAATGATCTGCAAACTCTCTTTTATGATGAAGATGAA 900  
DB 866 ACAAATTAATTAAGTAAATGATCTGCAAACTCTCTTTTATGATGAAGATGAA 925  
QY 901 AATGTATGATTTCTGCAATGATGCTTCTGCAATGATGCTTCTGCAATGATGCTTCT 960  
DB 926 AATGTATGATTTCTGCAATGATGCTTCTGCAATGATGCTTCTGCAATGATGCTTCT 985  
QY 961 GTAGAAATTTTAAACAGAGTACAGATGATGATGATGATGATGATGATGATGATGATG 1020  
DB 986 GTAGAAATTTTAAACAGAGTACAGATGATGATGATGATGATGATGATGATGATGATG 1045  
QY 1021 AATGTATGATTTCTGCAATGATGCTTCTGCAATGATGCTTCTGCAATGATGCTTCT 1080  
DB 1046 AATGTATGATTTCTGCAATGATGCTTCTGCAATGATGCTTCTGCAATGATGCTTCT 1105  
QY 1081 AAGCATTTCTCAACTTCAAGCAGCAGTGTGAGTAACTGCTGCTGCTGCTGCTGCTG 1140  
DB 1106 AAGCATTTCTCAACTTCAAGCAGCAGTGTGAGTAACTGCTGCTGCTGCTGCTGCTG 1165  
QY 1141 ACTGTGAGAGAAAGAAATTAATGATGATGATGATGATGATGATGATGATGATGATG 1200  
DB 1166 ACTGTGAGAGAAAGAAATTAATGATGATGATGATGATGATGATGATGATGATGATG 1225  
QY 1201 TACTGAGCGGATCTTAAAGTGAATCAGAAATGAGAACTGCTGCAAAATTAATTA 1260  
DB 1226 TACTGAGCGGATCTTAAAGTGAATCAGAAATGAGAACTGCTGCAAAATTAATTA 1285  
QY 1261 ACTGTGAGAGAAAGAAATTAATGATGATGATGATGATGATGATGATGATGATGATG 1320  
DB 1286 ACTGTGAGAGAAAGAAATTAATGATGATGATGATGATGATGATGATGATGATGATG 1345  
QY 1321 CTACAGTGAAGGCTTATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1380  
DB 1346 CTACAGTGAAGGCTTATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1405  
QY 1381 AAGTCTCTAGTAAAGACATACATCACTAACTAAACAGAGATGAATATTAAGTGGGA 1440  
DB 1406 AAGTCTCTAGTAAAGACATACATCACTAACTAAACAGAGATGAATATTAAGTGGGA 1465  
QY 1441 TCGCTTTTGAAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1500  
DB 1466 TCGCTTTTGAAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1525  
QY 1501 GTATTCAGAGGACAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1560  
DB 1526 GTATTCAGAGGACAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1585  
QY 1561 CCAGAAATTTCTGGAATCCAAAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1620  
DB 1586 CCAGAAATTTCTGGAATCCAAAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1645  
QY 1621 GAAATTAATGATGATTTTAAATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1680  
DB 1646 GAAATTAATGATGATTTTAAATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1705  
QY 1681 CTATTTGGAAGTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1740  
DB 1706 CTATTTGGAAGTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1765

Jan 20 08:47:21

**THIS PAGE BLANK (USPTO)**



1741 ACACAGCCTGATCCATAGTGGGATTTAGCTGTGACAAAGCTGTAATCTGATGAT 1800  
1766 ACACAGCCTGATCCATAGTGGGATTTAGCTGTGACAAAGCTGTAATCTGATGAT 1825  
1801 GCTCTTAATGATATTAATGAGAAATGATGATGATGATGATGATGATGATGATGAT 1860  
1826 GCTCTTAATGATATTAATGAGAAATGATGATGATGATGATGATGATGATGATGAT 1885  
1861 TATTTATTTAGGCAATTTGATGATGATGATGATGATGATGATGATGATGATGAT 1920  
1886 TATTTATTTAGGCAATTTGATGATGATGATGATGATGATGATGATGATGATGAT 1945  
1921 TTGACAGATGCAACCTCAGAGAGATTTATTTAGATGATGATGATGATGATGATGAT 1980  
1946 TTGACAGATGCAACCTCAGAGAGATTTATTTAGATGATGATGATGATGATGATGAT 2005  
1981 GCTGAGAGTTTATGAGAGAAATGAGAGACATATTTAGATGATGATGATGATGATGAT 2040  
2006 GCTGAGAGTTTATGAGAGAAATGAGAGACATATTTAGATGATGATGATGATGATGAT 2065  
2041 GGTAGAGTCCATATGTCGAAAGCATTTTCCAGAGACTTGGATGATGATGATGATGAT 2100  
2066 GGTAGAGTCCATATGTCGAAAGCATTTTCCAGAGACTTGGATGATGATGATGATGAT 2125  
2101 ATGGGTTACAGATTTTACAGAAATTTGAGATGATGATGATGATGATGATGATGAT 2160  
2126 ATGGGTTACAGATTTTACAGAAATTTGAGATGATGATGATGATGATGATGATGAT 2185  
2161 GTGGCTACGTGTTTGTGATCTCTGAGAGCTGGGCTTGAATGATGATGATGATGATGAT 2220  
2186 GTGGCTACGTGTTTGTGATCTCTGAGAGCTGGGCTTGAATGATGATGATGATGATGAT 2245  
2221 GAGCTCCAGGCTTCCAAACATTTTCAATTTTGAATCTTCCATCTGATGATGATGAT 2280  
2246 GAGCTCCAGGCTTCCAAACATTTTCAATTTTGAATCTTCCATCTGATGATGATGAT 2305  
2281 GGTGAGAAATTTGCTTGAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2340  
2306 GGTGAGAAATTTGCTTGAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2365  
2341 AAGGTAATCATTTGAGAAATGAGAAATTTGATGATGATGATGATGATGATGATGAT 2400  
2366 AAGGTAATCATTTGAGAAATGAGAAATTTGATGATGATGATGATGATGATGATGAT 2425  
2401 GCCACAGGCAACAGAGACCTTCTGATGATGATGATGATGATGATGATGATGATGAT 2460  
2426 GCCACAGGCAACAGAGACCTTCTGATGATGATGATGATGATGATGATGATGATGAT 2485  
2461 CCCATCAGGCAACAGATCTTGGAGAAATTTCTATCACTGATGATGATGATGATGATGAT 2520  
2486 CCCATCAGGCAACAGATCTTGGAGAAATTTCTATCACTGATGATGATGATGATGATGAT 2545  
2521 GCTTCTGATGCTGTGACCCAGATGATTTTATGATGATGATGATGATGATGATGATGAT 2580  
2546 GCTTCTGATGCTGTGACCCAGATGATTTTATGATGATGATGATGATGATGATGATGAT 2605  
2581 TCACATTCATCTTATTTAGACTTGAATGATGATGATGATGATGATGATGATGATGAT 2640  
2606 TCACATTCATCTTATTTAGACTTGAATGATGATGATGATGATGATGATGATGATGAT 2665  
2641 AGTTTCTCATTTCTCTTAATGATGATGATGATGATGATGATGATGATGATGATGAT 2700  
2666 AGTTTCTCATTTCTCTTAATGATGATGATGATGATGATGATGATGATGATGATGAT 2725  
2701 GAGATGTTCTTGTGCTTCCATCAATGCTTATGCTGATGATGATGATGATGATGATGAT 2760  
2726 GAGATGTTCTTGTGCTTCCATCAATGCTTATGCTGATGATGATGATGATGATGATGAT 2785  
2761 TGTGTTGAAACAGAAATGATTAATTTTGTCTCAATTTTATTTATTTATTTATTTAT 2820  
2786 TGTGTTGAAACAGAAATGATTAATTTTGTCTCAATTTTATTTATTTATTTATTTAT 2845

2821 AAAAAGAAACATGACGATTAATTTTGAAGAAAGCTCTTCAATTTATGAGCAAGT 2880  
2846 AAAAAGAAACATGACGATTAATTTTGAAGAAAGCTCTTCAATTTATGAGCAAGT 2905  
2881 TACAGAGAAACCTTCTCTATCAGAGGAGAAATGATGATGATGATGATGATGATGAT 2940  
2906 TACAGAGAAACCTTCTCTATCAGAGGAGAAATGATGATGATGATGATGATGATGAT 2965  
2941 GACCTTCTGAGACACTTGGGTTGATGATGATGATGATGATGATGATGATGATGATGAT 3000  
2966 GACCTTCTGAGACACTTGGGTTGATGATGATGATGATGATGATGATGATGATGATGAT 3025  
3001 CCTTACATGATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3060  
3026 CCTTACATGATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3085  
3061 CAGAAATCCAGGAGATTTTGGGATCAGAGAGAGATGATGATGATGATGATGATGAT 3120  
3086 CAGAAATCCAGGAGATTTTGGGATCAGAGAGAGATGATGATGATGATGATGATGAT 3145  
3121 GGCATTAAGTCCAGATTAACCTTACAGCTTATTTGATGATGATGATGATGATGATGAT 3180  
3146 GGCATTAAGTCCAGATTAACCTTACAGCTTATTTGATGATGATGATGATGATGATGAT 3205  
3181 AAGTATCAGCTTAACATTTGATGATGATGATGATGATGATGATGATGATGATGAT 3240  
3206 AAGTATCAGCTTAACATTTGATGATGATGATGATGATGATGATGATGATGATGAT 3265  
3241 AGAGAAATTTGAGCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3300  
3266 AGAGAAATTTGAGCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3325  
3301 AGTCTTAAGGAGAGAGAGCTTGAATGATGATGATGATGATGATGATGATGATGATGAT 3360  
3326 AGTCTTAAGGAGAGAGAGCTTGAATGATGATGATGATGATGATGATGATGATGATGAT 3385  
3361 ATGCAATTTCTGGGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3420  
3386 ATGCAATTTCTGGGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3445  
3421 GATATGAGATGAGAGCTTGAATGATGATGATGATGATGATGATGATGATGATGATGAT 3480  
3446 GATATGAGATGAGAGCTTGAATGATGATGATGATGATGATGATGATGATGATGATGAT 3505  
3481 GGAATCCCAATTTATGAGGCTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3540  
3506 GGAATCCCAATTTATGAGGCTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3565  
3541 ACTCAGATTAACCTGCTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3600  
3566 ACTCAGATTAACCTGCTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3625  
3601 GAAAGGCAAAATTAATCAAGTGAACGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3660  
3626 GAAAGGCAAAATTAATCAAGTGAACGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3685  
3661 GTACAGCCATGAGAGATTAATTTCCGAAATGATGATGATGATGATGATGATGATGAT 3720  
3686 GTACAGCCATGAGAGATTAATTTCCGAAATGATGATGATGATGATGATGATGATGAT 3745  
3721 AATGTTGATTAATGATGAGAGCTTGGGTTCTTGAAGAGAGAGAGAGAGAGAGAGAG 3780  
3746 AATGTTGATTAATGATGAGAGCTTGGGTTCTTGAAGAGAGAGAGAGAGAGAGAGAG 3805  
3781 CAAGAGCTTGAATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3840  
3806 CAAGAGCTTGAATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3865  
3841 GATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3900  
3866 GATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3925  
3901 GTTAACTTAATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3960

**THIS PAGE BLANK (USPTO)**



**THIS PAGE BLANK (USPTO)**

score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

## SUMMARIES

(without alignments)  
11702.894 Million cell updates/sec

11702.894 Million cell updates/sec

.....ttatggaactttggctgtga 4287

5777422

Post-processing: Minimum Match 0%

Listing first 45 summaries

GenEmbl: \*

1.1. gb\_ba:\*  
 2.2. gb\_hng:\*  
 3.3. gb\_in:\*  
 4.4. gb\_om:\*  
 5.5. gb\_ov:\*  
 6.6. gb\_pat:\*  
 7.7. gb\_ph:\*  
 8.8. gb\_pl:\*  
 9.9. gb\_pr:\*  
 10.10. gb\_ro:\*  
 11.11. gb\_sts:\*  
 12.12. gb\_ay:\*  
 13.13. gb\_un:\*  
 14.14. gb\_v1:\*  
 15.15. em\_ba:\*  
 16.16. em\_fun:\*  
 17.17. em\_hum:\*  
 18.18. em\_in:\*  
 19.19. em\_mu:\*  
 20.20. em\_om:\*  
 21.21. em\_or:\*  
 22.22. em\_ov:\*  
 23.23. em\_pat:\*  
 24.24. em\_ph:\*  
 25.25. em\_pl:\*  
 26.26. em\_ro:\*  
 27.27. em\_sts:\*  
 28.28. em\_un:\*  
 29.29. em\_v1:\*  
 30.30. em\_hng\_hum:\*  
 31.31. em\_hng\_in:\*  
 32.32. em\_hng\_other:\*  
 33.33. em\_hng\_mus:\*  
 34.34. em\_hng\_pln:\*  
 35.35. em\_hng\_rad:\*  
 36.36. em\_hng\_mam:\*  
 37.37. em\_hng\_vrt:\*  
 38.38. em\_ay:\*  
 39.39. em\_hng\_hum:\*  
 40.40. em\_hng\_mus:\*  
 41.41. em\_hng\_other:\*

Pred. No. is the number of results predicted by chance to have a

## SUMMARIES

Result	Query	
No.	Score	Match Length DB ID Description
-----		

1	43224.1	98.5	4761	6	AX534934	AX534934 Sequence
2	43224.4	98.5	5883	6	AT410459	AT410459 Homo sapi
3	43224.4	98.5	5895	6	AX534938	AX534938 Sequence
4	43222.8	98.5	4761	6	AX534936	AX534936 Sequence
5	43222.8	98.5	5895	6	AX534940	AX534940 Sequence
6	43221.4	98.5	4335	6	AX537610	AX537610 Sequence
7	43219.8	98.4	4335	6	AX537612	AX537612 Sequence
8	3187.4	74.4	3535	6	AX534942	AX534942 Sequence
9	3185.8	74.3	3335	6	AX534944	AX534944 Sequence
10	2709.2	63.2	4735	10	AY083458	AY083458 Mus muscu
11	2709.2	63.2	5644	10	BC052443	BC052443 Mus muscu
12	2448.2	57.1	2938	9	HSM805581	AT834478 Homo sapi
13	1969.4	45.9	2273	9	AK095888	AK095888 Homo sapi
14	730.8	17.0	1300	6	AX083146	AX083146 Sequence
15	230	5.4	163577	9	AX0590428	AX083146 Sequence
16	230	5.4	176444	2	AC012408	AX590428 Human DNA
17	219.6	5.1	5603	3	CIN431688	AC012408 Homo sapi
18	203.8	4.8	4678	3	DB3196	AY431688 Clona int
19	182	4.2	251250	2	AC097023	DB3196 Limulus sp.
20	177.4	4.1	91419	9	AX591480	AC097023 Rattus no
21	175	4.1	149597	2	AC034271	AX591480 Homo sapi
22	174	4.1	4527	6	AX054954	AC034271 Homo sapi
23	174	4.1	4560	6	AX054954	AX054953 Sequence
24	171.6	4.0	2629	3	AY118302	AY118302 Drosophill
25	167.6	3.9	4715	5	GGOVOM	AY118302 Drosophill
26	167	3.9	4578	5	AY122084	AY122084 Drosophill
27	167	3.9	4579	3	DME269539	AY122084 Drosophill
28	160	3.7	144589	2	AC112668	AY269539 Drosophill
29	160	3.7	191258	2	AC118254	AC118254 Mus muscu
30	155.4	3.6	4620	10	RNA113	AC118254 Mus muscu
31	154.8	3.6	176144	2	AC012408	AX5284 Rat mRNA fo
32	154.2	3.6	4636	10	FATX1113A	AC012408 Homo sapi
33	152.2	3.6	4220	10	DME269538	U03552 Rat plasma
34	151.2	3.5	4575	10	GPIMSPB	AY269538 Drosophill
35	149.8	3.5	2501	10	RATX113	DB4339 Cavia porce
36	146.4	3.4	4950	5	XELXENDO	U03524 Rat alpha-1
37	143.8	3.4	123013	3	AC092399	LG3513 Xenopus lae
38	143.8	3.4	164414	3	AC092399	AC092399 Drosophill
39	143.8	3.4	168503	3	AC115484	AC092399 Drosophill
40	143.8	3.4	276599	3	AE003649	AC115484 Drosophill
41	143.8	3.4	331718	3	DROSADH08	AE003649 Drosophill
42	140.8	3.3	431254	2	AC118254	AE003414 Drosophill
43	136	3.2	4623	10	BC058299	AC118254 Mus muscu
44	136	3.2	4672	10	BC051037	BC058299 Mus muscu
45	134.8	3.1	2689	5	AB026130	BC051037 Mus muscu
					AB026130	Cyprinus

## ALIGNMENTS

RESULT 1				
AX534934				
LOCUS	AX534934	4761 bp	DNA	linear
DEFINITION	Sequence 1 from Patent WO20070696.			
ACCESSION	AX534934			
VERSION	AX534934.1	GI:25261474		
KEYWORDS				
SOURCE	Homo sapiens (human)			
ORGANISM	Homo sapiens			
Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.				
REFERENCE	1			
AUTHORS	Schub, A. and Sutherland, R.D.			
TITLE	Cd109 nucleic acid molecules polypeptides and methods of use			
JOURNAL	Patent: WO 02070696-A 1 12-SEP-2002;			

**THIS PAGE BLANK (USPTO)**



gene  
5' UTR  
CDS

misc\_feature

3' UTR

478

4451..5883

/gene="CD109"

/gene="CD109"

/gene="CD109"

/db\_xref="taxon:9606"

/clone="X1"

1..5883

/gene="CD109"

1..112

/gene="CD109"

113..4450

/gene="CD109"

/note="associated with the Gov alloantigen system"

/codon\_start=1

/product="CD109"

/protein\_id="AAL64159.1"

/db\_xref="GI:19071209"

/translation="WGQPELITTAHLICVCTAALAVARPPRIYATGPIIRGGNVTI  
GVELLEHCPSQVTKAKELKTASMLTVSLVAEGVFEGSKRTLLIPSLPLNSADEIY  
ELRYATGTODEILFSNSTRLSFETKRLISVFLOTKALVKKQEVKFRVLTLSDFRPY  
KTSNTGILKIDPKSNLIQOMLSQCSODLGIKRTPLSHPIIGDVAIQVNDQTYOS  
FOSEVYTLKPREVTLOTPLCYCSNKLNGITATKATYGGKPGKVDVTLFELSPMGK  
KANITTKPKINGSANFSPNDEMKNMSNGLSYLLDSGPEILITVYESYGI  
SRNVSTVFFKQHDYITIEFPDYYLYKSLNPTATVKTATDGNQNTLEBRNNVIT  
VTQNYHSLWMSGNSGNQKMAVOAKINTVQSGTFKLEPILDESELQKAYLGS  
KTSNAHSLFKSPSKTYIQLKTRDENIKVSGPELVSGNKRKLKSLSTNVSRGOLVA  
VGSQNSTWFLSLTPENSTPCKACVIVYLEDGEIISDVKIPVOLVFNRKILXYVSK  
VKGQNSTWFLSLTPENSTPCKACVIVYLEDGEIISDVKIPVOLVFNRKILXYVSK  
KAESSEKSLRISVTQPSIVGIVAVDGSVNLMA NSDIIMENVHHEILMTVGYLG  
MPPNSAPFORCGIMVLTDA NLTQKDYIDGVYD NAEYARFMEENGHIVDIDFSGS  
SPHRKAPPEPTMIWLDTMGKGRVIOEPEVTVPDSTISVAGFVISEDGLATTPV  
ELQAFQPEFLINLPIYSITVIRGEERALETTIPVILKDADEVKYTEKSDPOTILMNE  
INATCHOQTLLVPSDEGATVLPFIRPTHGEIPITVTLSPASDAVQOMITLVKGEI  
EKSYQSILDLTDNRKLTSLTSLFSPEPNTVTSERVOITAGIDVDSINGLASL  
IRMPGSGEONNINFAPNVYILDYTLKKKOLDINKERKALSPRQOYQKELLVQREDG  
SPSANGVDPSGTSWLSAFVLRCPLEADPYIDIQNVILHRTYTWLKGKSGSEFWD  
GRVISHSELQGNKSPVPTLVAVTSLILGKRYQPNIDQESLHFLSESRGISDNYT  
IALITLTVLSVSGSPKAKALNLTWRABOEGMOPWSESLSDSMORSRLDI EVAAL  
YALLSHFQFQTSBGIPTIRWLSRONSISGRASSTODPTVALKALSEFALNMTERTN  
IQVATYGPSSPSPKVFLLIDTHNRLLQKEDLVAQMAVNIANSNGFALICOLNVVYN  
VVASGSSRRRSRIQMOEAFDLVAVAKENDLDNHDVLANCTSPSGGRGMALEVNVL  
LSGFVPSPEARLISLSTVKVEYEDHGLNLYIDSVETQCVNIPAVNRKFSVTDQAS  
VSIIVYPERQAVRSYNSSEVXLSSCDLSDVQCGRCPEDGASGSHHSVYIPIFCFK  
LLXPEEIMLT"

**THIS PAGE BLANK (USPTO)**

Db	413	ATTATGAGCTACGTGTAAACCGGACGTACCGAGATGAGATTTTATCTCTAATGATAC	472
Qy	361	CGGTATTCATTTGAGACCAAGGAAATATCTGTCTTCAATCAACACAGCAAGGCTTATAC	420
Db	473	CGGTATTCATTTGAGACCAAGGAAATATCTGTCTTCAATCAACACAGCAAGGCTTATAC	532
Qy	421	AAGCAAAGCAAGATGAAGTTCGTCATTTGATCACTCTCTCAGATTTTAAAGCTTAC	480
Db	533	AAGCAAAGCAAGATGAAGTTCGTCATTTGATCACTCTCTCAGATTTTAAAGCTTAC	592
Qy	481	AAAACCTCTTTAAACATTCCTCAATTAGGACCCCAATCAAAATTTGATCCAGAGTGTTG	540
Db	593	AAAACCTCTTTAAACATTCCTCAATTAGGACCCCAATCAAAATTTGATCCAGAGTGTTG	652
Qy	541	TCACAACAAAGATCTTGGAGTCATTTCCAAACTTTTACGTATCTTCCATCCATA	600
Db	653	TCACAACAAAGATCTTGGAGTCATTTCCAAACTTTTACGTATCTTCCATCCATA	712
Qy	601	CTTGGAGTCTGCTATTCAGGTCAGAGTATGACAGACATATTTATCAATTCATTCAG	660
Db	713	CTTGGAGTCTGCTATTCAGGTCAGAGTATGACAGACATATTTATCAATTCATTCAG	772
Qy	661	GTTTCAGAAATATGATTAACAAAATTTGAGATCTTTGACAGACACCATTAATGTTCT	720
Db	773	GTTTCAGAAATATGATTAACAAAATTTGAGATCTTTGACAGACACCATTAATGTTCT	832
Qy	721	ATGATTTCTAAGCATTTTAATGTAACATCAGGGCAAGATATATGGAAGCAGTG	780
Db	833	ATGATTTCTAAGCATTTTAATGTAACATCAGGGCAAGATATATGGAAGCAGTG	892
Qy	781	AAAGGAGCGTAAAGCTTACATTTTACCTTATCTTTTGGGAAAGAGAAAATAT	840
Db	893	AAAGGAGCGTAAAGCTTACATTTTACCTTATCTTTTGGGAAAGAGAAAATAT	952
Qy	841	ACAAAACATTTAAGATTAATGATCTGCAAACTCTCTTTAATAGTAGAAGATGAA	900
Db	953	ACAAAACATTTAAGATTAATGATCTGCAAACTCTCTTTAATAGTAGAAGATGAA	1012
Qy	901	AATGTAATGATTTCTTCAATGAGCTTTCTGATACCTGATCTATCTTCCCTGACCA	960
Db	1013	AATGTAATGATTTCTTCAATGAGCTTTCTGATACCTGATCTATCTTCCCTGACCA	1072
Qy	961	GTAGAAATTTTAACACAGTGCAGAAATCAGTATACAGATTTTCAAGAATGTAGACT	1020
Db	1073	GTAGAAATTTTAACACAGTGCAGAAATCAGTATACAGATTTTCAAGAATGTAGACT	1132
Qy	1021	AATGTGTTCTTCAAGCAACATGATTAACATCATGAGTTTGTGATTAATACATGCTTGG	1080
Db	1133	AATGTGTTCTTCAAGCAACATGATTAACATCATGAGTTTGTGATTAATACATGCTTGG	1192
Qy	1081	AAGCATCTCTCAACTTACAGCCACTGTGAAAGTAACTGCTGTAATGCAACCACTG	1140
Db	1193	AAGCATCTCTCAACTTACAGCCACTGTGAAAGTAACTGCTGTAATGCAACCACTG	1252
Qy	1141	ACTTTTGAAGAAAGAAAGAAATATATGATGCTTAAACATGTAACACAGAGAACTATAC	1200
Db	1253	ACTTTTGAAGAAAGAAAGAAATATATGATGCTTAAACATGTAACACAGAGAACTATAC	1312
Qy	1201	TACTGAGACGGATCTTAACAGTGAAGAAATCAGAAAAATGAAAGCTGTTCAGAAATTAAT	1260
Db	1313	TACTGAGACGGATCTTAACAGTGAAGAAATCAGAAAAATGAAAGCTGTTCAGAAATTAAT	1372
Qy	1261	ACTGTCCCCCAAGTGAACCTTTTAAGATTAATTCCTCAATCTGAGAGATTCAGGTAG	1320
Db	1373	ACTGTCCCCCAAGTGAACCTTTTAAGATTAATTCCTCAATCTGAGAGATTCAGGTAG	1432
Qy	1321	CTAAGTTGAAGGCTTATTTCTTGGTATGTAAGAAATGACATGCAAGTTCAATATGCTGTT	1380
Db	1433	CTAAGTTGAAGGCTTATTTCTTGGTATGTAAGAAATGACATGCAAGTTCAATATGCTGTT	1492
Qy	1381	AAGTCTCTTGAAGCATATACATCACTAAAAACAAGATGAAATATATTAAGTGTGA	1440

Db	1493	AAGTCTCCTAGTAAGACATACATCCAACTAAAAA	CAAGAGATGAAAAATATAAGGTGGGA	1557
Qy	1441	TCGCCCTTTTGAGTTGGTGGTGTAGTGGCAACAA	CGAATGGAAGAGTTAAGCATATAGTGA	1500
Db	1553	TGCGCTTTTGAGTTGGTGGTGTAGTGGCAACAA	CGAATGGAAGAGTTAAGCATATAGTGA	1612
Qy	1501	GTAATCCAGGGGACAGTTGGTGGTGTAGGAAAA	CAAAATTCACATGTTCTCTTTAACA	1566
Db	1613	GTATCCAGGGGACAGTTGGTGGTGTAGGAAAA	CAAAATTCACATGTTCTCTTTAACA	1672
Qy	1561	CCAGAAATTCGTGGACCTCCAAAAGCCCTGTGTA	ATGTGCTATTTATTTGAAGATGATGGG	1620
Db	1673	CCAGAAATTCGTGGACCTCCAAAAGCCCTGTGTA	ATGTGCTATTTATTTGAAGATGATGGG	1732
Qy	1621	GAAATTAATAGTGAATGTTCTTAAAAATTCCTGT	CAAGCTGTGTTTTTAAATAATTAAG	1680
Db	1733	GAAATTAATAGTGAATGTTCTTAAAAATTCCTGT	CAAGCTGTGTTTTTAAATAATTAAG	1792
Qy	1661	CTATATTGAGTAAAGTGAAGCTGAACCATGTAGA	AAAGTCTCTCTTAGATCTCTGTG	1748
Db	1793	CTATATTGAGTAAAGTGAAGCTGAACCATGTAGA	AAAGTCTCTCTTAGATCTCTGTG	1852
Qy	1741	ACAAGCTGTACTCATATGTTGGATTTGTAGCTGT	GAACAAAGTGTGAATCTGATGAAT	1800
Db	1853	ACAAGCTGTACTCATATGTTGGATTTGTAGCTGT	GAACAAAGTGTGAATCTGATGAAT	1912
Qy	1801	GCCTCTAATGATATTAACAATGAAAAATGTGTGC	ATGATGGAACCTTTATTAACAACAGA	1866
Db	1913	GCCTCTAATGATATTAACAATGAAAAATGTGTGC	ATGATGGAACCTTTATTAACAACAGA	1972
Qy	1861	TATTAATTTAGGACATGTTCAATGAATTCCTTTG	CACTCTTTCAGAAATGTGACCTCGGGTA	1920
Db	1973	TATTAATTTAGGACATGTTCAATGAATTCCTTTG	CACTCTTTCAGAAATGTGACCTCGGGTA	2032
Qy	1921	TTGACAGATGAAACCTTCACAGAGATTATTTGAT	GTATGATATTCATGACTTTCTTTG	1980
Db	2033	TTGACAGATGAAACCTTCACAGAGATTATTTGAT	GTATGATATTCATGACTTTCTTTG	2092
Qy	1981	GCTAGAGGTTTATGAGGAAAAATGAAAGACATAT	TGTGAATATTCATGACTTTCTTTG	2040
Db	2093	GCTAGAGGTTTATGAGGAAAAATGAAAGACATAT	TGTGAATATTCATGACTTTCTTTG	2152
Qy	2041	GGTAGAGTCCACATGTCGGAAGACATTTTCAGAGA	CTTGAGTTTGGCTTACAGACCAAC	2100
Db	2153	GGTAGAGTCCACATGTCGGAAGACATTTTCAGAGA	CTTGAGTTTGGCTTACAGACCAAC	2212
Qy	2101	ATGGGTTACAGGATTNACCAAGATTGGAAGTAA	CTGTACCTGATTTCTATCACTTCTGGG	2160
Db	2213	ATGGGTTACAGGATTNACCAAGATTGGAAGTAA	CTGTACCTGATTTCTATCACTTCTGGG	2272
Qy	2161	GTGGCTACCTGGTTTGTGATCTCTGAGAGACCTG	GGTCTTGGATTAACAACCTACAGTG	2220
Db	2273	GTGGCTACCTGGTTTGTGATCTCTGAGAGACCTG	GGTCTTGGATTAACAACCTACAGTG	2332
Qy	2221	GAGCTCCAGCCCTTCACACATTTTTCATTTTGTG	AATCTTCCATCTGTATTCAGA	2280
Db	2333	GAGCTCCAGCCCTTCACACATTTTTCATTTTGTG	AATCTTCCATCTGTATTCAGA	2392
Qy	2281	GGTAGAAGATTGCTTGGAAATTAACATATTAATTA	GAATGACCACTGAAGTT	2340
Db	2393	GGTAGAAGATTGCTTGGAAATTAACATATTAATTA	GAATGACCACTGAAGTT	2452
Qy	2341	AAGGTATATCATTTGAGAAAGTGCACAAATTTGAT	TATTCATATGACCTTCAAGTAATTAAT	2400
Db	2453	AAGGTATATCATTTGAGAAAGTGCACAAATTTGAT	TATTCATATGACCTTCAAGTAATTAAT	2512
Qy	2401	GCCACAGGCCACCAAGACAGCCCTCTGGTCTCCAG	TGAGAGATGGGGCAACTGTTCTTTT	2460
Db	2513	GCCACAGGCCACCAAGACAGCCCTCTGGTCTCCAG	TGAGAGATGGGGCAACTGTTCTTTT	2572
Qy	2461	CCCATCAGGCCCAACATCTTGAGAGAAATCTCTAT	CAACAGTCAAGCTTTTCAACCACT	2520
Db	2573	CCCATCAGGCCCAACATCTTGAGAGAAATCTCTAT	CAACAGTCAAGCTTTTCAACCACT	2632

**THIS PAGE BLANK (USPTO)**

2521 GCTTCTGATGCTGTGACCCAGATGATTTTAAAGGCTGAAGAAATAGAAATCATAT 2580  
 2633 GCTTCTGATGCTGTGACCCAGATGATTTTAAAGGCTGAAGAAATAGAAATCATAT 2692  
 2581 TCACAAATCATCTTATAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 2640  
 2693 TCACAAATCATCTTATAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 2752  
 2641 AGTTCTCATTTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2700  
 2753 AGTTCTCATTTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2812  
 2701 GGAGATGATTTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2760  
 2813 GGAGATGATTTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2872  
 2761 TGTGTGAAACAGACATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2820  
 2873 TGTGTGAAACAGACATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2932  
 2821 AAAAGAAACATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2880  
 2933 AAAAGAAACATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2992  
 2881 TACCAAGAGAACTTCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 2940  
 2993 TACCAAGAGAACTTCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 3052  
 2941 GACCCCTTGGAG 3000  
 3053 GACCCCTTGGAG 3112  
 3001 CCTTACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3060  
 3113 CCTTACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3172  
 3061 CAGAAATCCAG 3120  
 3173 CAGAAATCCAG 3232  
 3121 GGCAATTAAG 3180  
 3233 GGCAATTAAG 3292  
 3181 AAGTATCAGCTTAACTGATGATGATGATGATGATGATGATGATGATGATGATGAT 3240  
 3293 AAGTATCAGCTTAACTGATGATGATGATGATGATGATGATGATGATGATGATGAT 3352  
 3241 AAGGAAATTCAGACATTAATGATGATGATGATGATGATGATGATGATGATGATGAT 3300  
 3353 AAGGAAATTCAGACATTAATGATGATGATGATGATGATGATGATGATGATGATGAT 3412  
 3301 AGTCCCAAG 3360  
 3413 AGTCCCAAG 3472  
 3361 AAGCAATTCAG 3420  
 3473 AAGCAATTCAG 3532  
 3421 GATATGAAG 3480  
 3533 GATATGAAG 3592  
 3481 GATATGAAG 3540  
 3593 GATATGAAG 3652  
 3541 GATATGAAG 3600  
 3653 GATATGAAG 3712

3601 GAAAGACAAATATCCAG 3651  
 3713 GAAAGACAAATATCCAG 3772  
 3652 -----CTGCTGTGTGACAGCA 3669  
 3773 CTGATGACACAGACAG 3832  
 3670 ATGCAAGTAAATATTCAG 3729  
 3833 ATGCAAGTAAATATTCAG 3892  
 3730 TATATGAG 3789  
 3893 TATATGAG 3952  
 3790 TTTGATTTAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3849  
 3953 TTTGATTTAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4012  
 3850 GTGTGTACAGCTTTTGGGCGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3909  
 4013 GTGTGTACAGCTTTTGGGCGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4072  
 3910 TTAAGTGGCTTATGAG 3969  
 4073 TTAAGTGGCTTATGAG 4132  
 3970 GAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4029  
 4133 GAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4192  
 4030 GTTATATATTCCTGCTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4089  
 4193 GTTATATATTCCTGCTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4252  
 4090 ATATGAGATTAATCTATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4149  
 4253 ATATGAGATTAATCTATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4312  
 4150 CTGTCTCTCTGTGAG 4209  
 4313 CTGTCTCTCTGTGAG 4372  
 4210 TCAGGCTCCATCATCATCTTCACTGATGATGATGATGATGATGATGATGATGATGAT 4269  
 4373 TCAGGCTCCATCATCATCTTCACTGATGATGATGATGATGATGATGATGATGATGAT 4432  
 4270 ATGGAACCTTGGCTGTGA 4287  
 4433 ATGGAACCTTGGCTGTGA 4450

RESULT 3  
 AX534938 5895 bp DNA linear PAT 22-NOV-2002  
 LOCUS Sequence 5 from Patent WO02070696.  
 DEFINITION AX534938  
 ACCESSION AX534938  
 VERSION AX534938.1 GI:25261482  
 KEYWORDS  
 SOURCE  
 ORGANISM  
 Homo sapiens (human)  
 Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homini; Homo.  
 REFERENCE  
 1 Schub A. and Sutherland R.D.  
 C409 nucleic acid molecules polypeptides and methods of use  
 Patent: WO 02070696-A 5 12-SEP-2002;  
 Schub, Andre (Schub, Sutherland, Robert D. (CA)  
 Location/Qualifiers  
 1..5895  
 /organism="Homo sapiens"  
 /mol\_type="genomic DNA"

Save Result

**THIS PAGE BLANK (USPTO)**



RESULT 6  
AX537610  
LOCUS Sequence 1 from Patent WO02070738. 4335 bp DNA linear PAT 23-NOV-2002  
AX537610  
ACCESSION AX537610  
VERSION AX537610.1 GI:25269531  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.  
REFERENCE 1  
AUTHORS Schuh, A. and Ouwehand, W.  
TITLE Diagnosis and treatment of blood disorders  
JOURNAL Patent: WO 02070738-A 1 12-SEP-2002;  
Schuh, *Ante* *Ante*  
FEATURES  
source  
CDS  
1..4335  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
1..4335  
/note="unnamed protein product"  
/codon\_start=1  
/protein\_id="CAD57496.1"  
/db\_xref="GI:25269532"  
/translation="MOGPELITAAHLCTCTAALAVAPGRPLVTAAGIIRRGAVTI  
GRLBHCPSQVTVKALKTANLTVSVLEAGVBEKSPKLTLPSPILNADRIY  
ELAVTRTODEILFNSSTRILSFETKISVITQDKALYRKQVETRIYLPEDPDY  
KTSNLILIKDKPKNLIOQMSDGLKISFTQSLPILGWSILOVNDITYOS  
FQSEVLPKPEVTLQTPKXSNMKNLITAKTYGKPVAGDVLTPLEPFWGK  
KQITTKFKINGSANFSPNDEKMNWDSNGISEYDLSPPVEILLTVTSVGI  
SRVSTVFPKQNDYIIEFDYTVLPKPLNTPTRPVTPADQNLLEERNNVYT  
VTQNTVEYSGNSGKQKMAVOKINIVPQSTPKIEPVPALDSEELKAPLGS  
KSSMAVSLFKPSPKTYIOLKTRDENIKVSPFELVSGNRLEKELSYMVVSGOLVA  
VGKQNTMSELTSPNSWPKACVIVYIIEDEGEIISVLKIPVLVEKNIKLYMSV  
KABSEKVSRLISTPQDSIVGIIVADKSVLMAASNDIMENNVHLELYNTYLG  
MKNPSFAVQECGMVLTDALTDQYIDGVNDAEYARFENEHGHIVDHSLDS  
SPHAKHEPETWIMLDITMGRIYQEPPEVPSITSMVAVGFIIEKSDKPIIMTNS  
ELQAFQPFPIPLNPSVIRGEERLEITINYLKDAVEVYIIEKSDKPIIMTNS  
INATGHOQLTVPSDEGATVLEPIRPHLGEIPIYTVTLSTPAADATQMLVVAEG  
EKYSOSILDLTDNRLOSTLKLSPSPNIVGSEVQITAGDVPSEINLAEL  
IRMPYGCQKMINIAPNIVYIDYLTQKQKLDLKEALFPMQGYORELYQREDS  
SFAAGNDPDSGWTLSAFLRCFLEADYIDIDONVLRITYMLKQKNGEWD  
GRVHSELQGGKSPVYITAYIVTSILGRKQPNIDVOESIHLESEFSGISDNYT  
LALITVALSVSPKAKELNMLTRABQGMOPWSESLSMSQPRSLDLEVA  
VALLSHFLOFQVSEGIPIRMLSRQNSLGFASDQDTVALKALSEFALAMTERTN  
IOVTYGPSSPVKFLDTHNRLLQTLLELAVOPMAVNSFSGRSGALQANVYN  
VAGSGSPRRBSIIONEAFDLVAKENDDLNVDLAVNCTSPSGRSGALQANVYN  
LSGFWPSRLAISLSTVKKVEYDHGKLNLYDSVNETPCVNI PAVRNKVSNTDAS  
VSLVYDPEPRQKVSINSEVTLSDLSQDVQGCPCDASGSHHSVIFICFK  
LAFYFELML"

COUNT 1296 a 837 c 922 g 1280 t

98.5%; Score 4221.4; DB 6; Length 4335;  
Similarity 98.8%; P-Val. No. 0;  
Conservative 0; Mismatches 1; Indels 51; Gaps 1;

1 ATGCAAGGCGCCACCGGCTCTGACCGCGCCACCTCTCTGCTGTCACCGCGCGCTG 60  
1 ATGCAAGGCGCCACCGGCTCTGACCGCGCCACCTCTCTGCTGTCACCGCGCGCTG 60  
61 GCCGCGCTCCCGGCGCTCGGTTCTGCTGTCACCGCGCCACCTCTCTGCTGTCACCGCGCGCTG 120  
61 GCCGCGCTCCCGGCGCTCGGTTCTGCTGTCACCGCGCCACCTCTCTGCTGTCACCGCGCGCTG 120  
121 AATGCTATTTGGGCTGAGCTCTTGTGACACACGCGCGCTTCAAGGTGACTGTAAGGCG 180  
121 AATGCTATTTGGGCTGAGCTCTTGTGACACACGCGCGCTTCAAGGTGACTGTAAGGCG 180  
181 GAGCTGCTCAACACGACATCAACCTCACTGCTCTGCTGTCACCGCGCCACCTCTGTAAGGCG 240  
181 GAGCTGCTCAACACGACATCAACCTCACTGCTCTGCTGTCACCGCGCCACCTCTGTAAGGCG 240  
241 GAAAGGCTCTTTAAGACACTTACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 300  
241 GAAAGGCTCTTTAAGACACTTACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 300  
301 ATTATGAGCTACGTGTAACCGGCTGACCGGCTGACCGGCTGACCGGCTGACCGGCTGACCGGCTG 360  
301 ATTATGAGCTACGTGTAACCGGCTGACCGGCTGACCGGCTGACCGGCTGACCGGCTGACCGGCTG 360  
361 CGCTTATCTTTGAGACCAAGAAATATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 420  
361 CGCTTATCTTTGAGACCAAGAAATATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 420  
421 AAGCCAAAGCAAGAGTGAAGTTCGATGTTACCTCTTCTGCACTCTTCTGCACTTTTAAAGCTTAC 480  
421 AAGCCAAAGCAAGAGTGAAGTTCGATGTTACCTCTTCTGCACTCTTCTGCACTTTTAAAGCTTAC 480  
481 AAAACCTCTTTAAACATTTCTGATTAAGAACCCCAATCAATTTGATGCAACAGTGTG 540  
481 AAAACCTCTTTAAACATTTCTGATTAAGAACCCCAATCAATTTGATGCAACAGTGTG 540  
481 AAAACCTCTTTAAACATTTCTGATTAAGAACCCCAATCAATTTGATGCAACAGTGTG 540  
541 TCACAAACAAAGATCTTGAGTCAATTTCAAACTTTTCAAGCTTATCTTCCATCAATTA 600  
541 TCACAAACAAAGATCTTGAGTCAATTTCAAACTTTTCAAGCTTATCTTCCATCAATTA 600  
601 CTGCTGAGCTGCTGATTAAGCTTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAG 660  
601 CTGCTGAGCTGCTGATTAAGCTTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAG 660  
661 GTTTCAGATATGATTAACCAAAATTTGAAGTCACTTTGACAGACACATTAATTTGTTCT 720  
661 GTTTCAGATATGATTAACCAAAATTTGAAGTCACTTTGACAGACACATTAATTTGTTCT 720  
721 ATGAACTTCAAGATTTAATGATGATCAACGCGCAAGATATCATATGGAAGCAAGT 780  
721 ATGAACTTCAAGATTTAATGATGATCAACGCGCAAGATATCATATGGAAGCAAGT 780  
781 AAGGAGAGTAAAGCTTACATTTTACCTTTTGGGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAG 840  
781 AAGGAGAGTAAAGCTTACATTTTACCTTTTGGGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAG 840  
841 ACAAAAACATTTAAGATTAATGATGATGCAAACTTCTTTTAAATGATGATGATGATGATGATGATG 900  
841 ACAAAAACATTTAAGATTAATGATGATGCAAACTTCTTTTAAATGATGATGATGATGATGATGATG 900  
901 AATGTAATGATTTCTTAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 960  
901 AATGTAATGATTTCTTAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 960  
961 GTAGAAATTTTAAACAGTGAACAGTGAACAGTGAACAGTGAACAGTGAACAGTGAACAGTGAACAGT 1020  
961 GTAGAAATTTTAAACAGTGAACAGTGAACAGTGAACAGTGAACAGTGAACAGTGAACAGTGAACAGT 1020  
1021 AATGCTTCTTCAACAGTGAACAGTGAACAGTGAACAGTGAACAGTGAACAGTGAACAGTGAACAGT 1080  
1021 AATGCTTCTTCAACAGTGAACAGTGAACAGTGAACAGTGAACAGTGAACAGTGAACAGTGAACAGT 1080  
1081 AAGCATCTCTCAACAGTGAACAGTGAACAGTGAACAGTGAACAGTGAACAGTGAACAGTGAACAGT 1140

**THIS PAGE BLANK (USPTO)**

Db	2161	GTGGCTACTGCTGTTTGTGATCTCTGAGACCTGGGCTTGGACCTAACACTACTCCAGTG	2220
QY	2221	GAGCTCCAGCCCTTCCAACTTTTTCATTTTTTTTGAATCTTCCACTCTGTATTCAGA	2280
Db	2221	GAGCTCCAGCCCTTCCAACTTTTTCATTTTTTTTGAATCTTCCACTCTGTATTCAGA	2280
QY	2281	GGTGAAGATTTGCTTTGGAAATTAATATATTCATTAATTTTGAAGAATGCCACTGAGGTT	2340
Db	2281	GGTGAAGATTTGCTTTGGAAATTAATATATTCATTAATTTTGAAGAATGCCACTGAGGTT	2340
QY	2341	AAGGTATCATTTGAGAAAAGTGACAAATTTGATATTTCTATGACTCTCAAGTGAATTAAT	2400
Db	2341	AAGGTATCATTTGAGAAAAGTGACAAATTTGATATTTCTATGACTCTCAATGAATTAAT	2400
QY	2401	GCCACAGGCCACCAACAACCTTCTGGTTCCTCAGTGAAGATGGGGCAACTGTTCCTTTT	2460
Db	2401	GCCACAGGCCACCAACAACCTTCTGGTTCCTCAGTGAAGATGGGGCAACTGTTCCTTTT	2460
QY	2461	CCCATCAGGCCCAACACATCTGGGAGAAATTCCTATCAAGTCAAGCTCTTTCACCACT	2520
Db	2461	CCCATCAGGCCCAACACATCTGGGAGAAATTCCTATCAAGTCAAGCTCTTTCACCACT	2520
QY	2521	GCTTCTGATGCTGTGCACCCAGATGATTTTGTAGTAAAGGCTGAAGAAATGAATATCAT	2580
Db	2521	GCTTCTGATGCTGTGCACCCAGATGATTTTGTAGTAAAGGCTGAAGAAATGAATATCAT	2580
QY	2581	TCACAATCCACTTATTTATGACTTTGACTGAACAATAGGCTTACAGAGTACCCTGAAAATTG	2640
Db	2581	TCACAATCCACTTATTTATGACTTTGACTGAACAATAGGCTTACAGAGTACCCTGAAAATTG	2640
QY	2641	AGTTTCAATTTCTCCTTAATPACAGTGCCTGGGAGTGAAGAAGTTCAATCACTGCAATT	2700
Db	2641	AGTTTCAATTTCTCCTTAATPACAGTGCCTGGGAGTGAAGAAGTTCAATCACTGCAATT	2700
QY	2701	GGAGATGTTCTTGGTCCCTTCATCAATGCTTACCTCACTGATTCGGATGCTTATGGC	2760
Db	2701	GGAGATGTTCTTGGTCCCTTCATCAATGCTTACCTCACTGATTCGGATGCTTATGGC	2760
QY	2761	TGTGTGCAACAGAAACATGATTAATTTGCTCCAAAATTTACATTTTGGATTTATCTGACT	2820
Db	2761	TGTGTGCAACAGAAACATGATTAATTTGCTCCAAAATTTACATTTTGGATTTATCTGACT	2820
QY	2821	AAAAAGAAACAACGACAGATATTTTGAAGAAAAGACCTCTTCATTTATGAGGCAAGT	2880
Db	2821	AAAAAGAAACAACGACAGATATTTTGAAGAAAAGACCTCTTCATTTATGAGGCAAGT	2880
QY	2881	TACACAGAGAACTTCTATTCAGAGGGAAGATGGCTCTTCAGTGGCTTTTGGAAATTAAT	2940
Db	2881	TACACAGAGAACTTCTATTCAGAGGGAAGATGGCTCTTCAGTGGCTTTTGGAAATTAAT	2940
QY	2941	GACCTTCTGGAGACACTTGGTTTGAGCTTTTGTTTTAAAGATGTTTCTTGAAGCCGAT	3000
Db	2941	GACCTTCTGGAGACACTTGGTTTGAGCTTTTGTTTTAAAGATGTTTCTTGAAGCCGAT	3000
QY	3001	CCCTACATGATATTTGATCAGAAATGTGTACACAGAAACATCACTTGGCTTAAAGACAT	3060
Db	3001	CCCTACATGATATTTGATCAGAAATGTGTGTACACAGAAACATCACTTGGCTTAAAGACAT	3060
QY	3061	CAGAAATCCACGATGAATTTTGGGATCCAGAGAAGGATTCATATGTGAGCTTCAAGGT	3120
Db	3061	CAGAAATCCACGATGAATTTTGGGATCCAGAGAAGGATTCATATGTGAGCTTCAAGGT	3120
QY	3121	GGCAATTAAGTCCAGTACACTTACAGCTATATTTGTAATCTTCTCCTGGGATTAAGA	3180
Db	3121	GGCAATTAAGTCCAGTACACTTACAGCTATATTTGTAATCTTCTCCTGGGATTAAGA	3180
QY	3181	AAGTATGAGCCTAACATGATCTGTGAAGAGTCTATCAATTTTGGAGTCTGAATTCAGT	3240
Db	3181	AAGTATGAGCCTAACATGATGATGTGAAGAGTCTATCAATTTTGGAGTCTGAATTCAGT	3240
QY	3241	AGAGAAATTTGACAATTTACTCTAGCCCTTAACTTATGCAATTTGCAATTCATGGGG	3300
Db	3241	AGAGAAATTTGACAATTTACTCTAGCCCTTAACTTATGCAATTTGCAATTCATGGGG	3300

**THIS PAGE BLANK (USPTO)**



PLANK (USPTO)



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 16, 2004, 12:36:15 ; Search time 1017 Seconds  
(without alignments)  
11379.047 Million cell updates/sec

Title: US-10-020-095-3

Perfect score: 4287  
Sequence: 1 atgcagggccacgcgcctc.....ttatggaacttgctgtga 4287

Scoring table: IDENTITY NUC  
Gapop 10.0, Gapext 1.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 510512

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database :

1: /SIDSI/gcgdata/geneeq/geneeqn-emb1/NA1980.DAT:\*  
2: /SIDSI/gcgdata/geneeq/geneeqn-emb1/NA1981.DAT:\*  
3: /SIDSI/gcgdata/geneeq/geneeqn-emb1/NA1982.DAT:\*  
4: /SIDSI/gcgdata/geneeq/geneeqn-emb1/NA1983.DAT:\*  
5: /SIDSI/gcgdata/geneeq/geneeqn-emb1/NA1984.DAT:\*  
6: /SIDSI/gcgdata/geneeq/geneeqn-emb1/NA1985.DAT:\*  
7: /SIDSI/gcgdata/geneeq/geneeqn-emb1/NA1986.DAT:\*  
8: /SIDSI/gcgdata/geneeq/geneeqn-emb1/NA1987.DAT:\*  
9: /SIDSI/gcgdata/geneeq/geneeqn-emb1/NA1988.DAT:\*  
10: /SIDSI/gcgdata/geneeq/geneeqn-emb1/NA1989.DAT:\*  
11: /SIDSI/gcgdata/geneeq/geneeqn-emb1/NA1990.DAT:\*  
12: /SIDSI/gcgdata/geneeq/geneeqn-emb1/NA1991.DAT:\*  
13: /SIDSI/gcgdata/geneeq/geneeqn-emb1/NA1992.DAT:\*  
14: /SIDSI/gcgdata/geneeq/geneeqn-emb1/NA1993.DAT:\*  
15: /SIDSI/gcgdata/geneeq/geneeqn-emb1/NA1994.DAT:\*  
16: /SIDSI/gcgdata/geneeq/geneeqn-emb1/NA1995.DAT:\*  
17: /SIDSI/gcgdata/geneeq/geneeqn-emb1/NA1996.DAT:\*  
18: /SIDSI/gcgdata/geneeq/geneeqn-emb1/NA1997.DAT:\*  
19: /SIDSI/gcgdata/geneeq/geneeqn-emb1/NA1998.DAT:\*  
20: /SIDSI/gcgdata/geneeq/geneeqn-emb1/NA1999.DAT:\*  
21: /SIDSI/gcgdata/geneeq/geneeqn-emb1/NA2000.DAT:\*  
22: /SIDSI/gcgdata/geneeq/geneeqn-emb1/NA2001.DAT:\*  
23: /SIDSI/gcgdata/geneeq/geneeqn-emb1/NA2002.DAT:\*  
24: /SIDSI/gcgdata/geneeq/geneeqn-emb1/NA2003.DAT:\*  
25: /SIDSI/gcgdata/geneeq/geneeqn-emb1/NA2003.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	4281.2	99.9	4369	25	AAD49434 Human r150 gene #1
2	4224.4	98.5	4761	24	ABO79964 Human CD109 K1 pro
3	4224.4	98.5	5882	25	AAD49440 Human blood cell s
4	4224.4	98.5	5883	24	ABO79966 Human CD109 K1-H7
5	4223.4	98.5	5883	25	AAD49435 Human r150 DNA #1
6	4223.8	98.5	4761	24	ABO79965 Human CD109 K1 var
7	4222.8	98.5	5895	24	ABO79967 Human CD109 K1-H7
8	4221.4	98.5	4335	24	AAL49815 Human platelet all

9	4219.8	98.4	4335	24	AAL49816 Human platelet all
10	4219.6	98.4	4473	21	AAA62010 Hydrophobic domain
11	4216.6	98.4	4335	21	AAA60199 Hydrophobic domain
12	4140.2	96.6	4146	25	AAD49436 Human blood cell s
13	4082.4	95.2	4197	25	AAD49437 Human r150 DNA #2
14	3187.4	74.4	3535	24	ABO79968 Human CD109 K15 pr
15	3185.8	74.3	3535	24	ABO79969 Human CD109 K15 va
16	1017	23.7	1448	24	ABL90478 Human polynucleoti
17	975.4	22.8	1459	22	ABA08825 Human secreted pro
18	812.2	18.9	2403	24	ABV77330 Alpha-1 protease
19	730.8	17.0	1300	22	AAAF1735 Human secretase and
20	541.6	12.6	875	20	AAAX0449 Human gene express
21	525	11.6	821	25	ABX08843 Angiogenesis-associ
22	497.6	11.6	744	20	AAZ15521 Human gene express
23	331.6	7.7	690	20	AAZ13539 DNA encoding novel
24	300	7.0	300	20	AAZ13539 Human gene express
25	272	6.3	354	25	ABX47702 Bovine EST associat
26	175	4.1	744	20	AAZ16725 Human gene express
27	174	4.1	4527	22	AAZ09511 C. elegans alpha-2
28	174	4.1	4527	22	AAZ09511 Drosophila melanog
29	172.6	4.0	4560	22	AAZ09511 Drosophila melanog
30	171.6	4.0	2285	23	ABL20611 Drosophila melanog
31	152.2	3.6	3612	23	ABL19705 Drosophila melanog
32	143.8	3.4	6324	23	ABL20610 Drosophila melanog
33	127.2	3.0	5611	23	ABL20610 Drosophila melanog
34	127.2	3.0	6979	23	AAZ12152 Drosophila melanog
35	126.8	3.0	4615	23	AAZ12152 DNA encoding novel
36	126.8	3.0	4615	24	ABN92132 Gene #3730 used to
37	126.8	3.0	4629	23	AAZ92192 Human ORF39 P
38	126.8	3.0	5816	21	AAZ74484 DNA encoding novel
39	125.2	2.9	4487	24	ABK2035 Human alpha-2-macr
40	125.2	2.9	4488	24	ABK2035 Human pregnancy zo
41	120.8	2.8	4660	24	ABO93897 Human platelet all
42	120.6	2.8	2608	24	AAL49817 DNA encoding human
43	119.8	2.8	4501	25	ABX70464 Human cDNA SEQ ID
44	119.4	2.8	1140	22	ABA06596 Human polynucleoti
45	119.4	2.8	1140	22	ABA06869 Human polynucleoti

## ALIGNMENTS

RESULT 1	AAD49434	AAAD49434 standard; DNA; 4369 BP.
XX	XX	XX
AC	AAAD49434;	
XX	XX	XX
DT	24-MAR-2003	(first entry)
XX	XX	XX
DE	Human r150 gene #1.	
XX	XX	XX
KW	Human; GPI-anchored TGF-beta1 binding protein; r150 protein; cancer; glycosylphosphatidy inositol; transforming growth factor-beta1;	
KW	therapy; gene; ds.	
XX	XX	XX
OS	Homo sapiens.	
XX	XX	XX
PH	Key	Location/Qualifiers
FT	CDS	26..4312
FT	FT	/+tag= a
FT	FT	/product= "Human r150 protein #1"
FT	FT	/transl_except= (pos:2132..2134, aa:Xaa)
FT	FT	/note= "Xaa corresponds to Ser, Tyr"
XX	XX	XX
XX	XX	XX
PN	WO200285942-A2	
XX	XX	XX
XX	XX	XX
PD	31-OCT-2002.	
XX	XX	XX
XX	XX	XX
PF	24-APR-2002; 2002WO-CR00560.	
XX	XX	XX
XX	XX	XX
PR	24-APR-2001; 2001US-285713P.	
PR	14-FEB-2002; 2002US-356163P.	

**THIS PAGE BLANK (USPTO)**

XX (UTMC-) UNIV MCGILL.  
PA Phillip A, Tam B;  
PI WPI: 2003-093100/08.  
DR P-PSDB; AAE32012.  
XX Novel transforming growth factor (TGF)-beta 1 binding reagent which  
PT comprises r150 protein which acts as accessory receptor of TGF-beta,  
PT useful for negatively modulating TGF-beta activity, and thus for  
PT treating cancer  
XX  
PS Claim 7; Page 91-93; 127p; English.  
XX The invention relates to novel transforming growth factor (TGF)-beta1  
CC binding reagent which comprises a glycosylphosphatidyl inositol (GPI)-  
CC anchored TGF-beta binding protein referred to as r150 which acts as  
CC an accessory receptor of TGF-beta. The invention is used for negatively  
CC modulating TGF-beta activity, and thus for treating conditions  
CC characterized by overproduction of TGF-beta, such as cancer. Antisense  
CC molecules of the invention are used for increasing TGF-beta availability  
CC and increase graft success. The present sequence is a gene encoding  
CC human r150 protein.  
XX  
SQ Sequence 4369 BP; 1301 A; 842 C; 932 G; 1293 T; 1 other;  
Query Match 99.9%; Score 4281.2; DB 25; Length 4369;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 4283; Conservative 0; Mismatches 4; Indels 0; Gaps 0;  
QY 1 ATGAGAGGCGCCGCTCTGACCGCGCCGCACTCTCTGCTGTCAGCCGCGCTG 60  
DB 26 ATGAGAGGCGCCGCTCTGACCGCGCCGCACTCTCTGCTGTCAGCCGCGCTG 85  
QY 61 GCCGTGGCTCCCGGCGCTCTGCTGTCAGCCGCGCCGCACTCTCTGCTGTCAGCCGCGCTG 120  
DB 86 GCCGTGGCTCCCGGCGCTCTGCTGTCAGCCGCGCCGCACTCTCTGCTGTCAGCCGCGCTG 145  
QY 121 AATGTGACTATTGGGGTGAAGCTTCTGGAACATGCGCTTCAAGTGAAGTGAAGGCG 180  
DB 146 AATGTGACTATTGGGGTGAAGCTTCTGGAACATGCGCTTCAAGTGAAGTGAAGGCG 205  
QY 181 GAGCTGCTCAAGACAGCATCAACCTCACTCTCTGCTGTCAGCCGCGCTG 240  
DB 206 GAGCTGCTCAAGACAGCATCAACCTCACTCTCTGCTGTCAGCCGCGCTG 265  
QY 241 GAAAGAGGCTCTTTAAGACACTTACTCTTCATCACTCACTCGAAGAGAGAGGCTT 300  
DB 266 GAAAGAGGCTCTTTAAGACACTTACTCTTCATCACTCACTCGAAGAGAGAGGCTT 325  
QY 301 ATTATGAGCTACGTGAACCGAGCGTACCAGATGATTTTATTCTTAATAGTACC 360  
DB 326 ATTATGAGCTACGTGAACCGAGCGTACCAGATGATTTTATTCTTAATAGTACC 385  
QY 361 CGCTTATCATTTGAGACCAAGAAATATCTGCTTCAATCAACAGACAGGCTTATAC 420  
DB 386 CGCTTATCATTTGAGACCAAGAAATATCTGCTTCAATCAACAGACAGGCTTATAC 445  
QY 421 AAGCCAAAGCAAGAGTGAAGTTCCGATGTTTACACTCTCTCAAGATTTTAAGCCTTAC 480  
DB 446 AAGCCAAAGCAAGAGTGAAGTTCCGATGTTTACACTCTCTCAAGATTTTAAGCCTTAC 505  
QY 481 AAAACCTCTTTAAACATTTCTATTAAGAGCCCAATCAATTTGATCAACAGTGTG 540  
DB 506 AAAACCTCTTTAAACATTTCTATTAAGAGCCCAATCAATTTGATCAACAGTGTG 565  
QY 541 TCACAAAGAGTATCTTGAAGTCAATTTCAAACTTTTCACTATCTTCCATCAATA 600  
DB 566 TCACAAAGAGTATCTTGAAGTCAATTTCAAACTTTTCACTATCTTCCATCAATA 625  
QY 601 CTGTGATGCTGTATATCAATTTCAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 660

DB 626 CTGTGATGCTGTATATCAATTTCAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 685  
QY 661 GTTTCAGAAATATGATTAATTAACAAATTTTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 720  
DB 686 GTTTCAGAAATATGATTAATTAACAAATTTTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 745  
QY 721 ATGAATTTCAAGATTTTAAATGATTAACATCAACGCAAAAGTATCATATGGAAGCCAGTG 780  
DB 746 ATGAATTTCAAGATTTTAAATGATTAACATCAACGCAAAAGTATCATATGGAAGCCAGTG 805  
QY 781 AAGGAGAGCTAACGCTTACATTTTACCTTTATCTTTGGGAAAGAAATATTT 840  
DB 806 AAGGAGAGCTAACGCTTACATTTTACCTTTATCTTTGGGAAAGAAATATTT 865  
QY 841 ACAAACATTTAAGATTAATGATTAATGATTAATGATTAATGATTAATGATTAATGATTAAT 900  
DB 866 ACAAACATTTAAGATTAATGATTAATGATTAATGATTAATGATTAATGATTAATGATTAAT 925  
QY 901 AATGTAATGATTTCTTCAATGATTAATGATTAATGATTAATGATTAATGATTAATGATTAAT 960  
DB 926 AATGTAATGATTTCTTCAATGATTAATGATTAATGATTAATGATTAATGATTAATGATTAAT 985  
QY 961 GTTGAATTTTAAACAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1020  
DB 986 GTTGAATTTTAAACAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1045  
QY 1021 AATGTTCTTCAAGACATGATTAATGATTAATGATTAATGATTAATGATTAATGATTAATGATTAAT 1080  
DB 1046 AATGTTCTTCAAGACATGATTAATGATTAATGATTAATGATTAATGATTAATGATTAATGATTAAT 1105  
QY 1081 AAGCCATTTCTCAACTTCAACAGCTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1140  
DB 1106 AAGCCATTTCTCAACTTCAACAGCTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1165  
QY 1141 ACTCTTGAAGAAAGAAATTAATGATTAATGATTAATGATTAATGATTAATGATTAATGATTAAT 1200  
DB 1166 ACTCTTGAAGAAAGAAATTAATGATTAATGATTAATGATTAATGATTAATGATTAATGATTAAT 1225  
QY 1201 TACTGAGCGGATCTTCAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1260  
DB 1226 TACTGAGCGGATCTTCAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1285  
QY 1261 ACTGTCCTTCAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1320  
DB 1286 ACTGTCCTTCAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1345  
QY 1321 CTACAGTTGAAGGCTTATTTCTTGTATTAATGATTAATGATTAATGATTAATGATTAATGATTAAT 1380  
DB 1346 CTACAGTTGAAGGCTTATTTCTTGTATTAATGATTAATGATTAATGATTAATGATTAATGATTAAT 1405  
QY 1381 AAGTCTCTAGTGAAGATTAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAAT 1440  
DB 1406 AAGTCTCTAGTGAAGATTAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAAT 1465  
QY 1441 TCGCTTTTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1500  
DB 1466 TCGCTTTTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1525  
QY 1501 GTATCAAGGGGACAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1560  
DB 1526 GTATCAAGGGGACAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1585  
QY 1561 CCAGAAATTTCTTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1620  
DB 1586 CCAGAAATTTCTTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1645  
QY 1621 GAAATTTAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1680  
DB 1646 GAAATTTAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1705  
QY 1681 CTATATTGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1740  
DB 1706 CTATATTGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1765

**THIS PAGE BLANK (USPTO)**

QY 1741 ACAAGCCTGATCCATAGTGGATGAGTGTGACCAAAAGTGAATCTGAT 1800  
 DB 1766 ACACAGCTGATCCATAGTGGATGAGTGTGACCAAAAGTGAATCTGAT 1825  
 QY 1801 GCTCTTAATTAATTAATGAAGAAATGAGTGTGACCAAAAGTGAATCTGAT 1860  
 DB 1826 GCTCTTAATTAATTAATGAAGAAATGAGTGTGACCAAAAGTGAATCTGAT 1885  
 QY 1861 TATTATTAGGCAATTAATGAAGAAATGAGTGTGACCAAAAGTGAATCTGAT 1920  
 DB 1886 TATTATTAGGCAATTAATGAAGAAATGAGTGTGACCAAAAGTGAATCTGAT 1945  
 QY 1921 TTGACAGATGCAAACTCAGAAAGATTAATGAGTGTGACCAAAAGTGAATCTGAT 1980  
 DB 1946 TTGACAGATGCAAACTCAGAAAGATTAATGAGTGTGACCAAAAGTGAATCTGAT 2005  
 QY 1981 GCTGAGAGTTTAATGAAGAAATGAGTGTGACCAAAAGTGAATCTGAT 2040  
 DB 2006 GCTGAGAGTTTAATGAAGAAATGAGTGTGACCAAAAGTGAATCTGAT 2065  
 QY 2041 GGTAGAGTCCACATGTCGAAAGATTTCCAGAGCTTGAGTGTGACCAAAAGTGAATCTGAT 2100  
 DB 2066 GGTAGAGTCCACATGTCGAAAGATTTCCAGAGCTTGAGTGTGACCAAAAGTGAATCTGAT 2125  
 QY 2101 ATGGGTTACAGATTTACCAAGATTTGAGTGTGACCAAAAGTGAATCTGAT 2160  
 DB 2126 ATGGGTTACAGATTTACCAAGATTTGAGTGTGACCAAAAGTGAATCTGAT 2185  
 QY 2161 GTGGCTAATGTTTGTGATCTGAGAGCTTGAGTGTGACCAAAAGTGAATCTGAT 2220  
 DB 2186 GTGGCTAATGTTTGTGATCTGAGAGCTTGAGTGTGACCAAAAGTGAATCTGAT 2245  
 QY 2221 GAGCTCAAGCCTTCCAAACATTTTCAATTTTGAATCTTGAGTGTGACCAAAAGTGAATCTGAT 2280  
 DB 2246 GAGCTCAAGCCTTCCAAACATTTTCAATTTTGAATCTTGAGTGTGACCAAAAGTGAATCTGAT 2305  
 QY 2281 GGTGAAGAAATTTGTTGAAATTAATTAATGAAGTGTGACCAAAAGTGAATCTGAT 2340  
 DB 2306 GGTGAAGAAATTTGTTGAAATTAATTAATGAAGTGTGACCAAAAGTGAATCTGAT 2365  
 QY 2341 AAGGTATATGAGAAAGTGAATTAATTAATGAAGTGTGACCAAAAGTGAATCTGAT 2400  
 DB 2366 AAGGTATATGAGAAAGTGAATTAATTAATGAAGTGTGACCAAAAGTGAATCTGAT 2425  
 QY 2401 GGCACAGGCAACAGAGACCTTCTGAGTGTGACCAAAAGTGAATCTGAT 2460  
 DB 2426 GGCACAGGCAACAGAGACCTTCTGAGTGTGACCAAAAGTGAATCTGAT 2485  
 QY 2461 CCCATCAGGCAACATCTGAGAGAAATCTATCAAGTGTGACCAAAAGTGAATCTGAT 2520  
 DB 2486 CCCATCAGGCAACATCTGAGAGAAATCTATCAAGTGTGACCAAAAGTGAATCTGAT 2545  
 QY 2521 GCTTCTGATGCTGACCAAGATTTTAATGAAGTGTGACCAAAAGTGAATCTGAT 2580  
 DB 2546 GCTTCTGATGCTGACCAAGATTTTAATGAAGTGTGACCAAAAGTGAATCTGAT 2605  
 QY 2581 TCACATCCATCTTAATGAAGTGTGACCAAAAGTGAATCTGAT 2640  
 DB 2606 TCACATCCATCTTAATGAAGTGTGACCAAAAGTGAATCTGAT 2665  
 QY 2641 AGTTTCTCATTTCTCTTAATGAAGTGTGACCAAAAGTGAATCTGAT 2700  
 DB 2666 AGTTTCTCATTTCTCTTAATGAAGTGTGACCAAAAGTGAATCTGAT 2725  
 QY 2701 GAGAGATTTCTGCTCTTCAATGAAGTGTGACCAAAAGTGAATCTGAT 2760  
 DB 2726 GAGAGATTTCTGCTCTTCAATGAAGTGTGACCAAAAGTGAATCTGAT 2785  
 QY 2761 TGTGTGAACAGAAATGAATTTTGTGACCAAAAGTGAATCTGAT 2820  
 DB 2786 TGTGTGAACAGAAATGAATTTTGTGACCAAAAGTGAATCTGAT 2845

QY 2821 AAAAAGAAACATGACATTAATTTGAAGAAAGTCTTCTTCAATTAATGAAGAGGT 2880  
 DB 2846 AAAAAGAAACATGACATTAATTTGAAGAAAGTCTTCTTCAATTAATGAAGAGGT 2905  
 QY 2881 TACCAGAGAACTTCTTAATGAAGAGAGTCTTCTTCAATTAATGAAGAGGT 2940  
 DB 2906 TACCAGAGAACTTCTTAATGAAGAGAGTCTTCTTCAATTAATGAAGAGGT 2965  
 QY 2941 GACCTTCTGAGAGCACTTGTGAGTGTGACCAAAAGTGAATCTGAT 3000  
 DB 2966 GACCTTCTGAGAGCACTTGTGAGTGTGACCAAAAGTGAATCTGAT 3025  
 QY 3001 CTTTACATGATTTTATGAGTGTGACCAAAAGTGAATCTGAT 3060  
 DB 3026 CTTTACATGATTTTATGAGTGTGACCAAAAGTGAATCTGAT 3085  
 QY 3061 CAGAAATCCAAAGGTGAATTTTGAAGTGTGACCAAAAGTGAATCTGAT 3120  
 DB 3086 CAGAAATCCAAAGGTGAATTTTGAAGTGTGACCAAAAGTGAATCTGAT 3145  
 QY 3121 GGCATTAATAAGTCCAGTAACTTACAGCTTAATTTGAATCTTCTGAGATTAAG 3180  
 DB 3146 GGCATTAATAAGTCCAGTAACTTACAGCTTAATTTGAATCTTCTGAGATTAAG 3205  
 QY 3181 AAGTATCAGCTTAATGATGAGTGTGACCAAAAGTGAATCTGAT 3240  
 DB 3206 AAGTATCAGCTTAATGATGAGTGTGACCAAAAGTGAATCTGAT 3265  
 QY 3241 AAGAAATTTGAGCAATTAATGATGAGTGTGACCAAAAGTGAATCTGAT 3300  
 DB 3266 AAGAAATTTGAGCAATTAATGATGAGTGTGACCAAAAGTGAATCTGAT 3325  
 QY 3301 AGTCTTAATAAGCAAGAGTGTGACCAAAAGTGAATCTGAT 3360  
 DB 3326 AGTCTTAATAAGCAAGAGTGTGACCAAAAGTGAATCTGAT 3385  
 QY 3361 ATGCAATTTCTGAGTGTGACCAAAAGTGAATCTGAT 3420  
 DB 3386 ATGCAATTTCTGAGTGTGACCAAAAGTGAATCTGAT 3445  
 QY 3421 GATATTGAAGTGTGACCAAGTGTGACCAAAAGTGAATCTGAT 3480  
 DB 3446 GATATTGAAGTGTGACCAAGTGTGACCAAAAGTGAATCTGAT 3505  
 QY 3481 GGAATCCCAATTAATGAGTGTGACCAAAAGTGAATCTGAT 3540  
 DB 3506 GGAATCCCAATTAATGAGTGTGACCAAAAGTGAATCTGAT 3565  
 QY 3541 ACTCAGATTAACAGTGTGCTTAAAGTGTGACCAAAAGTGAATCTGAT 3600  
 DB 3566 ACTCAGATTAACAGTGTGCTTAAAGTGTGACCAAAAGTGAATCTGAT 3625  
 QY 3601 GAAAGCAAAATATCAAGTGTGACCAAAAGTGAATCTGAT 3660  
 DB 3626 GAAAGCAAAATATCAAGTGTGACCAAAAGTGAATCTGAT 3685  
 QY 3661 GTACAGCAATGAGTGTGACCAAAAGTGAATCTGAT 3720  
 DB 3686 GTACAGCAATGAGTGTGACCAAAAGTGAATCTGAT 3745  
 QY 3721 AATGTTGATTAATGAGTGTGACCAAAAGTGAATCTGAT 3780  
 DB 3746 AATGTTGATTAATGAGTGTGACCAAAAGTGAATCTGAT 3805  
 QY 3781 CAGAAGCTTGAATTAATGAGTGTGACCAAAAGTGAATCTGAT 3840  
 DB 3806 CAGAAGCTTGAATTAATGAGTGTGACCAAAAGTGAATCTGAT 3865  
 QY 3841 GATTTGAATGATTAATGAGTGTGACCAAAAGTGAATCTGAT 3900  
 DB 3866 GATTTGAATGATTAATGAGTGTGACCAAAAGTGAATCTGAT 3925  
 QY 3901 GTTAACATTAATGAGTGTGACCAAAAGTGAATCTGAT 3960

**THIS PAGE BLANK (USPTO)**



Db	3926	GTAAACCTAATTAAAGGCGCTTATAGTGCGCTTCAGAAACAATTCTCTGAGCGACAAGTG	3987
Qy	3961	AAGAAAGTGAATATGATCATGAGAAACCTCAACGCTCTATTAGATTCTGTAAATGAAACC	4022
Db	3986	AAGAAAGTGAATATGATCATGAGAAACCTCAACGCTCTATTAGATTCTGTAAATGAAACC	4045
Qy	4021	CAGTTTGTGTTAATATTTCTGCTGTGAGAACTTTAAAGTTCAAAATACCAGAATGCT	4086
Db	4046	CAGTTTGTGTTAATATTTCTGCTGTGAGAACTTTAAAGTTCAAAATACCAGAATGCT	4105
Qy	4081	TCAGTGTCCATAGTGGATTAATCTATATGAGCGAAGAGACAGCGGCTGAGAAATTCAACTCT	4140
Db	4106	TCAGTGTCCATAGTGGATTAATCTATATGAGCGAAGAGACAGCGGCTGAGAAATTCAACTCT	4165
Qy	4141	GAAGTGAAGCTGTCTCTCTGTGACCTTTGACAGTAGTATGTCAGAGGGGTGCGGCTTTGGAG	4200
Db	4166	GAAGTGAAGCTGTCTCTCTGTGACCTTTGACAGTAGTATGTCAGAGGGGTGCGGCTTTGGAG	4225
Qy	4201	GATGAGAGCTTCAAGGCTCCCATCATCATCTCTTCAGTCATTTTTTATTTTCTGTTCAAGCTT	4260
Db	4226	GATGAGAGCTTCAAGGCTCCCATCATCATCTCTTCAGTCATTTTTTATTTTCTGTTCAAGCTT	4285
Qy	4261	CTGTACTTTATGAGAACTTTGGCTGTGA	4287
Db	4286	CTGTACTTTATGAGAACTTTGGCTGTGA	4312

RESULT 2
ABQ79964
ID ABQ79964 standard; cDNA; 4761 BP.
XY

DT 23-DEC-2002 (first entry)  
yy

Human CD109 K1 protein encoding cDNA.

cardiovascular; vasotropic; gene therapy; CD109 Kl; gene; ss.

OS **Homo sapiens**.

[illegible]

PD 12-SEP-2002.

07-MAR-2002; 2002WO-CA00292

PR 07-MAR-2001; 2001US-273814P

PA (SCHU/) SCHUH A.  
PA (SUTH/) SUTHERLAND R D.

Schuh A, Sutherland RD;

DR WPI; 2002-713450/77.  
DR P-PSDB; ABB82165.

p1 New CD109 nucleic acids and polypeptides, useful in gene therapy,  
 p1 particularly for treating strokes, myocardial infarctions, thrombosis  
 p1 chromocytopenia, autoimmune diseases, or organ or bone marrow  
 p1 transplantation -

Claim 1; Fig 1a; 156pp; English

The invention relates to isolated nucleic acid molecules encoding CD105

polypeptides. These nucleic acid molecules include the human cDNA sequences comprising CD109 KI, CD109 KI-H7, CD109 KI5 or their variants. CD109 is a novel member of the alpha2 macroglobulin (alpha2M)/C3, C4, C5 family of thioester-containing proteins. The CD109 polypeptides can be expressed by standard recombinant methodology. The CD109 nucleic acid, CD109 polypeptide or its mimetic, or the anti-CD109 antibody is useful for treating or detecting a disease or disorder, e.g. conditions associated with endothelial activation, platelet activation, activation of the coagulation or fibrinolytic systems, activation of T lymphocytes and of the complement system, quantitative or qualitative abnormalities of platelet function, increased or impaired platelet aggregation and activation, increased or impaired activation of the coagulation and/or fibrinolytic systems, or impaired or increased immune activation. These are also useful for treating cardiovascular disorders, stroke, myocardial infarction, thrombosis, embolism, peripheral vascular disease, thrombocytopenia, thrombocytopenia, autoimmune diseases, organ transplantation, or bone marrow transplantation. The CD109 nucleic acids or their antisense nucleotide sequence are useful in gene therapy for treating these conditions. The present sequence represents the human CD109 KI cDNA sequence.

Sequence 4761 BP; 1420 A; 912 C; 1016 G; 1413 T; 0 other;

Query Match	98.5%;	Score 4224.4;	DB 24;	Length 4761;
Best Local Similarity	99.0%;	Score 4224.4;	DB 24;	Length 4761;

Matches 4286; Conservative 0; Mismatches 1; Indels 51; Gaps 1;

1 ATGCAGGGCCACCGCTCCTGACCGCGCCCACTCTCTGCGTGTGCACCGCGGCTG 60

113 ATGCAGGGCCCAACGGCTCCTGACCGCGCCACCTCCTGTGGTGACCGCGCGCTG 172

61. GCCGTCCTCCGGGCTCGGTTCTGGTGACAGCCCCAGGGATCATCAGGCCCGAGGA 120

1/3 GCGTGGCTCCCGGCTCCGTTCTGGTGACAGCCCCAGGGATCATCAGGCCCGGAGGA 237

121 AAAGACACATGGGCTGGAGCTTCTGGACACACTGCCCTTACACAGGTGACTGTGAAGCC 180

22 223 AACGACAAAGGAGCTCTGGACACTGCCCTTCACAGGTGACTGTGAAGCG 297

[illegible]

-----

[illegible][illegible][illegible]

.....

[illegible][illegible][illegible][illegible][illegible][illegible]

Db 653 TCACACAAAGTGATCTTGGAGTCA TTTCCAAA CTTTTCAGCTATCTTCCCATCCAAATA 712

601 CTTGGTGA CTGGTCTATTCAAGTTCAGTGAATGACCAAGACATATTTATCAATCATTTACG 660

Db 713 CTGGTACTGGTCTATTCAAGTTCAGTGATGACCAGACATATTCATCATTCAG 772

661 GTTCAGATATGTATTACCAAAATTGAGTGACTTTCAGACACCATTTATATTGTTCT 720

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

nucleic search, using sw model

January 16, 2004, 12:38:20 ; Search time 14986 Seconds  
(without alignments)  
11702.894 Million cell updates/sec

Sequence: 1 atcgagggccccacgcctcct.....ttatggaacttcgtgctgta 4287  
US-10-020-095-3  
4287

Scoring table:  
IDENTITY NUC  
Gapex 10.0, Gapext 1.0

Searched: 2888711 seqs, 2045481386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database: 1: gb\_ba:\*  
2: gb\_hcg:\*  
3: gb\_in:\*  
4: gb\_cm:\*  
5: gb\_ov:\*  
6: gb\_pat:\*  
7: gb\_ph:\*  
8: gb\_pl:\*  
9: gb\_dr:\*  
10: gb\_ro:\*  
11: gb\_str:\*  
12: gb\_by:\*  
13: gb\_un:\*  
14: gb\_vl:\*  
15: em\_ba:\*  
16: em\_fun:\*  
17: em\_hum:\*  
18: em\_in:\*  
19: em\_mu:\*  
20: em\_om:\*  
21: em\_or:\*  
22: em\_ov:\*  
23: em\_pat:\*  
24: em\_ph:\*  
25: em\_pl:\*  
26: em\_ro:\*  
27: em\_str:\*  
28: em\_un:\*  
29: em\_vl:\*  
30: em\_hcg\_hum:\*  
31: em\_hcg\_inv:\*  
32: em\_hcg\_other:\*  
33: em\_hcg\_mus:\*  
34: em\_hcg\_pla:\*  
35: em\_hcg\_rod:\*  
36: em\_hcg\_man:\*  
37: em\_hcg\_vtc:\*  
38: em\_ey:\*  
39: em\_hgo\_hum:\*  
40: em\_hgo\_mus:\*  
41: em\_hgo\_other:\*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	% Match	Query Length	DB ID	Description
1	4224.4	98.5	4761	AX534934	AX534934 Sequence
2	4224.4	98.5	5883	AP410459	AP410459 Homo sapi
3	4224.4	98.5	5895	AX534938	AX534938 Sequence
4	4222.8	98.5	4761	AX534936	AX534936 Sequence
5	4222.8	98.5	5895	AX534940	AX534940 Sequence
6	4221.8	98.5	4335	AX537610	AX537610 Sequence
7	4219.8	98.4	4335	AX537612	AX537612 Sequence
8	3187.4	74.3	3535	AX534942	AX534942 Sequence
9	3185.8	74.3	4735	AX534944	AX534944 Sequence
10	2709.2	63.2	5644	AY083458	AY083458 Mus muscu
11	2709.2	63.2	5644	BC052443	BC052443 Mus muscu
12	2448.2	57.1	2938	AK095888	AK095888 Homo sapi
13	1969.4	45.9	2273	AK095888	AK095888 Homo sapi
14	730.8	17.0	1300	AX083146	AX083146 Sequence
15	232	5.4	163577	AL590428	AL590428 Homo sapi
16	230	5.1	176144	AC012408	AC012408 Homo sapi
17	219.6	4.8	5603	CIN43168	CIN43168 Homo sapi
18	203.6	4.8	4678	D83196	D83196 Limulus sp.
19	182	4.2	251250	AC097023	AC097023 Rattus no
20	177.4	4.1	91419	AL591480	AL591480 Human DNA
21	175	4.1	149597	AC034271	AC034271 Homo sapi
22	174	4.1	4527	AX054954	AX054954 Sequence
23	174	4.1	4560	AX054954	AX054954 Sequence
24	171.6	4.0	2629	AY118302	AY118302 Drosophila
25	167.6	3.9	4715	GGOVOM	GGOVOM G.gallus mr
26	167	3.9	4578	AY122084	AY122084 Drosophila
27	167	3.9	4579	DME269539	DME269539 Drosophila
28	160	3.7	144589	AC112668	AC112668 Mus muscu
29	160	3.7	191258	AC118254	AC118254 Mus muscu
30	155.4	3.6	4620	RNA113	RNA113 Rat mRNA fo
31	154.8	3.6	176144	AC012408	AC012408 Homo sapi
32	154.2	3.6	4636	RAT113A	RAT113A Rat plasma
33	152.2	3.5	4220	DME269538	DME269538 Drosophila
34	151.2	3.5	4575	GPI MSPB	GPI MSPB Drosophila
35	149.8	3.5	2501	RAT113	RAT113 Rat plasma
36	146.4	3.4	4950	XELENDO	XELENDO Drosophila
37	143.8	3.4	122013	AC092399	AC092399 Drosophila
38	143.8	3.4	164414	AC092395	AC092395 Drosophila
39	143.8	3.4	168503	AC115484	AC115484 Drosophila
40	143.8	3.4	275699	AE003649	AE003649 Drosophila
41	143.8	3.4	337178	DROSADH08	DROSADH08 Drosophila
42	140.8	3.3	191258	AC118254	AC118254 Mus muscu
43	136	3.2	4623	BC036299	BC036299 Mus muscu
44	136	3.2	4672	BC051037	BC051037 Mus muscu
45	134.8	3.1	2689	AB026130	AB026130 Cypripus

## ALIGNMENTS

RESULT 1  
AX534934  
LOCUS AX534934 4761 bp DNA linear PAT 22-NOV-2002  
DEFINITION Sequence 1 from Patent WO02070696.  
ACCESSION AX534934  
VERSION AX534934.1 GI:25261474  
KEYWORDS  
SOURCE  
ORGANISM Homo sapiens (human)  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE  
1 Schuh, A. and Sutherland, R.D.  
TITLE Cdi09 nucleic acid molecules polypeptides and methods of use  
JOURNAL Patent: WO 02070696-A 1 12-SEP-2002;

**THIS PAGE BLANK (USPTO)**

Y	3670	ATGGCAGTTAAATTTTCCGCAAAATGGTTTTGATTTGCAATTTGGCAGTCAATGTGA	3729
b	3833	ATGGCAGTTAAATTTTCCGCAAAATGGTTTTGATTTGCAATTTGGCAGTCAATGTGA	3892
Y	3730	TATTAATGTGAAAGGCTTCTGGGCTCTTCTAGAAAGCAAGATCTATCCAAATCAAGAGCC	3789
b	3893	TATTAATGTGAAAGGCTTCTGGGCTCTTCTAGAAAGCAAGATCTATCCAAATCAAGAGCC	3952
Y	3790	TTTGATTTAGATGTTGCTGTGTAAAGAAATTAAGATGATCTCAATCAATGGATTTGAAT	3849
b	3953	TTTGATTTAGATGTTGCTGTGTAAAGAAATTAAGATGATCTCAATCATGTGAATTTGAAT	4012
Y	3850	GTGTGTACAAGCTTTTCGGGCGCGGAGTGAAGATGGATGGCTTTATGGAAGTTAACTTA	3909
b	4013	GTGTGTACAAGCTTTTCGGGCGCGGAGTGAAGATGGATGGCTTTATGGAAGTTAACTTA	4072
Y	3910	TTTAAGTGGCTTTATGATGCTTCAAGAACATTTCTCTGAGCAGACACTGAAAGAAAGTG	3969
b	4073	TTTAAGTGGCTTTATGATGCTTCAAGAACATTTCTCTGAGCAGACACTGAAAGAAAGTG	4132
Y	3970	GAATATGATCATGAGAAATCCAACTCTATTTAGATTCTGTAAATGAACCCAGTTTGT	4029
b	4133	GAATATGATCATGAGAAATCCAACTCTATTTAGATTCTGTAAATGAACCCAGTTTGT	4192
Y	4030	GTTAATATTTCTGCTGTGAGAAACTTTAAAGTTTCAAAATCCCAAGATGCTTCAGTGTCC	4089
b	4193	GTTAATATTTCTGCTGTGAGAAACTTTAAAGTTTCAAAATCCCAAGATGCTTCAGTGTCC	4252
Y	4090	ATATGATGATTAATCTATGAGCCAAAGAGCAAGCGGTGGAAGTTAAACAATCTGAAGTGAAG	4149
b	4253	ATATGATGATTAATCTATGAGCCAAAGAGCAAGCGGTGGAAGTTAAACAATCTGAAGTGAAG	4312
Y	4150	CTGTCTCTCTGTGACCTTTGCAAGTGAAGTCCAGGGCTCCGCTCTTGTGAGATGGAAGCT	4209
b	4313	CTGTCTCTCTGTGACCTTTGCAAGTGAAGTCCAGGGCTCCGCTCTTGTGAGATGGAAGCT	4372
Y	4210	TCAGGCTTCCCATCAATCACTCTTGAGTGAATTTTATCTGTTCAAGCTTCGTAATCTT	4269
b	4373	TCAGGCTTCCCATCAATCACTCTTGAGTGAATTTTATCTGTTCAAGCTTCGTAATCTT	4432
Y	4270	ATGGAACCTTTGGCTGTGA 4287	
b	4433	ATGGAACCTTTGGCTGTGA 4450	

RESULT 2  
APR410459

LOCUS	AF110459	5883 bp	mRNA	linear	PRI 02-MAR-2002
DEFINITION	Homo sapiens CD109 (CD109) mRNA, complete cds.				
ACCESSION	AF110459				
VERSION	AF110459.1	GI:19071208			
KEYWORDS					
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
REFERENCE	Lin, M., Sutherland, D.R., Horstfall, W., Totty, N., Yao, E., Nayer, R., Wu, X.F., and Schud, A.C.				
AUTHORS	Cell surface antigen CD109 is a novel member of the alpha(2) macroglobulin/C3, C4, C5 family of thioester-containing proteins				
TITLE	Blood 99 (5), 1683-1691 (2002)				
JOURNAL					
FEEDLINE	21683/742				

PUBMED 11661284  
 REFERENCE  
 AUTHORS  
 TITLE  
 JOURNAL  
 PAGES  
 Abstract  
 2 (bases 1 to 5893)  
 Lin, M., Sutherland, D. R., Horsfall, W., Totty, N., Yeo, E., Nayar, R.,  
 Wu, X.-F. and Schuh, A. C.  
 Direct Submission  
 Submitted (14-AUG-2001) Medicine, University of Toronto, 1 King's  
 College Circle, Room 7366, Toronto, Ontario M5S 1A6, Canada  
 Location/Qualifiers  
 1..5883  
 /organism="Homo sapiens"

```

gene
1..5883
/gene="CD109"
5'UTR
1..112
/gene="CD109"
CDS
113..4450
/gene="CD109"

```

BASE COUNT	misc_feature	3' UTR	misc_feature	BASE COUNT
1697 a	1697 a	1697 a	1697 a	1697 a
1173 c	1173 c	1173 c	1173 c	1173 c
1241 g	1241 g	1241 g	1241 g	1241 g
1772 t	1772 t	1772 t	1772 t	1772 t

Query Match	98.5%	Score 4224.4	DB 9	Length 5883
Best Local Similarity	98.8%	Pred. No. 0		
Matches 4286	Conservative	0	Indels 1	Gaps 1
QY	1	ATGCAAGGGCCACCGGCTCTGTACCGCGCCCACTCTCTGTGGTGTGACCCCGCGCTG	60	
Db	113	ATGCAAGGGCCACCGGCTCTGTACCGCGCCCACTCTCTGTGGTGTGACCCCGCGCTG	172	
QY	61	GCCGTGGCTCCCGGCGCTCGGTTTCTGTGTGACAGGCCCAAGGATCATCAGGCGCCGAGGA	120	
Db	173	GCCGTGGCTCCCGGCGCTCGGTTTCTGTGTGACAGGCCCAAGGATCATCAGGCGCCGAGGA	232	
QY	121	AATGTGATATTGGGGTGGAGGCTTGTGAAACACTGCGCTTCACAGGTGACTGTGAAGGGG	180	
Db	223	AATGTGATATTGGGGTGGAGGCTTGTGAAACACTGCGCTTCACAGGTGACTGTGAAGGGG	292	
QY	181	GAGCTGCTCAAGACAGCATCAAACTCACTGTCTCTGTCTTGTGGAAGCAGAAGAGCTTT	240	
Db	293	GAGCTGCTCAAGACAGCATCAAACTCACTGTCTCTGTCTTGTGGAAGCAGAAGAGCTTT	352	
QY	241	GAAAAAGGCTCTTTTAAGCACTTACTCTTCATCACTACTCTGTAACAGTGCAGATGAG	300	
Db	353	GAAAAAGGCTCTTTTAAGCACTTACTCTTCATCACTACTCTGTAACAGTGCAGATGAG	412	
QY	301	ATTATAGAGTACGTGTAAACCGAGCTAACCCAGATGAGATTTTATTCTCTAATAGTACC	360	

**THIS PAGE BLANK (USPTO)**



Db 413 ATTATAGCTAGCTGTAACGAGCGTACCGAGGATGAGTTTATCTCAATAGTACC 472  
Qy 361 CGCTTATCATTTTGAACCAAGAAATCTGTCTTCAATGCAACAGAGGCTTATAC 420  
Db 473 CGCTTATCATTTTGAACCAAGAAATCTGTCTTCAATGCAACAGAGGCTTATAC 532  
Qy 421 AAGCCAAAGCAAGAGTGAAGTTTCCGATTTGTAACCTCTTCTCAATTTTAAAGCTTAC 480  
Db 533 AAGCCAAAGCAAGAGTGAAGTTTCCGATTTGTAACCTCTTCTCAATTTTAAAGCTTAC 592  
Qy 481 AAAACCTCTTAAACATCTCTCAATTAAGGAGCCCAATCAAAATTTGATCCACAGTGTG 540  
Db 593 AAAACCTCTTAAACATCTCTCAATTAAGGAGCCCAATCAAAATTTGATCCACAGTGTG 652  
Qy 541 TCACAAACAAAGTATCTTGAGTCAATTTCCAAACCTTTCAAGCATCTTCCCAATTA 600  
Db 653 TCACAAACAAAGTATCTTGAGTCAATTTCCAAACCTTTCAAGCATCTTCCCAATTA 712  
Qy 601 CTGTGAGTCTGTCTATTCAGATTCAAGTGAATGACACACATATTAATCAATCAATTCAG 660  
Db 713 CTGTGAGTCTGTCTATTCAGATTCAAGTGAATGACACACATATTAATCAATCAATTCAG 772  
Qy 661 GTTTCAGAAATGATATTAACCAAAATTTGAAGTACATTTGCAACACCATTAATTTGTC 720  
Db 773 GTTTCAGAAATGATATTAACCAAAATTTGAAGTACATTTGCAACACCATTAATTTGTC 832  
Qy 721 ATGAATTTCAAGATTTTAATGTGATACATCAAGCAAGATTAATGAGAAAGCAAGT 780  
Db 833 ATGAATTTCAAGATTTTAATGTGATACATCAAGCAAGATTAATGAGAAAGCAAGT 892  
Qy 781 AAGAGAGAGTAAAGCTTACATTTTACCTTTATCTTTTGGGAAAGAAATATAT 840  
Db 893 AAGAGAGAGTAAAGCTTACATTTTACCTTTATCTTTTGGGAAAGAAATATAT 952  
Qy 841 ACAAACCACTTAAGTAAATGATATGATGCAAACTTCTTTTAAATGATGAAGATGAA 900  
Db 953 ACAAACCACTTAAGTAAATGATATGATGCAAACTTCTTTTAAATGATGAAGATGAA 1012  
Qy 901 AATGTATGATTTCTTCAAAATGATGATTTGTAATGCTGATCTATCTTCCCTGAGCA 960  
Db 1013 AATGTATGATTTCTTCAAAATGATGATTTGTAATGCTGATCTATCTTCCCTGAGCA 1072  
Qy 961 GTAGAAATTTTAAACACAGTACAGATCAAGTATTAAGAAATGTAAGCACT 1020  
Db 1073 GTAGAAATTTTAAACACAGTACAGATCAAGTATTAAGAAATGTAAGCACT 1132  
Qy 1021 AATGTGTTCTTCAACCAATGATTAATCAATGATTTTGTATTAATTAATGATGCTG 1080  
Db 1133 AATGTGTTCTTCAACCAATGATTAATCAATGATTTTGTATTAATTAATGATGCTG 1192  
Qy 1081 AAGCATCTCTCAACCTTCAACGACCTGTAAGGATTAATCTGAGTGAAGCAACACTG 1140  
Db 1193 AAGCATCTCTCAACCTTCAACGACCTGTAAGGATTAATCTGAGTGAAGCAACACTG 1252  
Qy 1141 ACTCTTGAAGAAAGAAATTAATGATGATTAATCAAGTACACAGAGAACTATAGTAG 1200  
Db 1253 ACTCTTGAAGAAAGAAATTAATGATGATTAATCAAGTACACAGAGAACTATAGTAG 1312  
Qy 1201 TACTGAGAGGATCTAACAGTGAATTAAGAAATGAAAGCTGTTCAAGAAATTAATTA 1260  
Db 1313 TACTGAGAGGATCTAACAGTGAATTAAGAAATGAAAGCTGTTCAAGAAATTAATTA 1372  
Qy 1261 ACTGTCCCAAAAGTGAACCTTTTAAGTGAATTTCCCAATCTGAGAGATTTCCAGTAG 1320  
Db 1373 ACTGTCCCAAAAGTGAACCTTTTAAGTGAATTTCCCAATCTGAGAGATTTCCAGTAG 1432  
Qy 1321 CTACAGTTGAAGGCTTATTTCTTGTGATGAATTAAGTACATGAGCTGATGCTGTTT 1380  
Db 1433 CTACAGTTGAAGGCTTATTTCTTGTGATGAATTAAGTACATGAGCTGATGCTGTTT 1492  
Qy 1381 AAGTCTCTGATGAACATCAATCAATTAACCAAGAGTGAATTAATTAAGGTTGGA 1440

Db 1493 AAGTCTCTGATGAACATCAATCAATTAACCAAGAGTGAATTAATTAAGGTTGGA 1552  
Qy 1441 TCGCCTTTTGAAGTGTGATGATGAGCAACAAAGATTAAGGATTAAGTGA 1500  
Db 1553 TCGCCTTTTGAAGTGTGATGATGAGCAACAAAGATTAAGGATTAAGTGA 1612  
Qy 1501 GTATCCAGGAGCAAGTGTGCTGTAGAGAAACAAATTTCAACAAATGTTCTTTAA 1560  
Db 1613 GTATCCAGGAGCAAGTGTGCTGTAGAGAAACAAATTTCAACAAATGTTCTTTAA 1672  
Qy 1561 CCAGAAATTTCTTGAAGCTCCAAAGGCTGTATATTTGTATTAATTAAGATGAG 1620  
Db 1673 CCAGAAATTTCTTGAAGCTCCAAAGGCTGTATATTTGTATTAATTAAGATGAG 1732  
Qy 1621 GAAATTAATGAAGTGTGATTAATAATCTGTGACCTGTTTAAATTAAGATGA 1680  
Db 1733 GAAATTAATGAAGTGTGATTAATAATCTGTGACCTGTTTAAATTAAGATGA 1792  
Qy 1681 CTATATGAGATTAAGTGAAGCTGAACCATCTGAGAAAGTCTCTTGAATCTGTG 1740  
Db 1793 CTATATGAGATTAAGTGAAGCTGAACCATCTGAGAAAGTCTCTTGAATCTGTG 1852  
Qy 1741 ACACAGCCTGATCCATGATGTTGGGATTTGATGCTGTGAACAAAGTGAATGATGA 1800  
Db 1853 ACACAGCCTGATCCATGATGTTGGGATTTGATGCTGTGAACAAAGTGAATGATGA 1912  
Qy 1801 GCTCTTAATGATTAATCAATGAGAAATGTGTCATGATGGAACCTTAATTAACAGGA 1860  
Db 1913 GCTCTTAATGATTAATCAATGAGAAATGTGTCATGATGGAACCTTAATTAACAGGA 1972  
Qy 1861 TATTAATTAAGCATTTTCAATGATTTCTTTGAGCTTTTCAGAAATGAGATCTGGGTA 1920  
Db 1973 TATTAATTAAGCATTTTCAATGATTTCTTTGAGCTTTTCAGAAATGAGATCTGGGTA 2032  
Qy 1921 TTGACAGATGCAACCTGACAGAGATTAATGATGCTGTTTATGAACATGACAGAT 1980  
Db 2033 TTGACAGATGCAACCTGACAGAGATTAATGATGCTGTTTATGAACATGACAGAT 2092  
Qy 1981 GCTGAGAGTTTATGAGAGAAATGAAGACATATGAGATTAATCATGATCTTCTTGG 2040  
Db 2093 GCTGAGAGTTTATGAGAGAAATGAAGACATATGAGATTAATCATGATCTTCTTGG 2152  
Qy 2041 GTGAGCATCTCAACATGCTCCGAAGCAATTTTCCAGAGCTTGATTTGCTAGACCAAC 2100  
Db 2153 GTGAGCATCTCAACATGCTCCGAAGCAATTTTCCAGAGCTTGATTTGCTAGACCAAC 2212  
Qy 2101 ATGGGTTACAGATTTTACCAAGATTTGAAGTAACTGAACCTGATTTTCACTCTTGG 2160  
Db 2213 ATGGGTTACAGATTTTACCAAGATTTGAAGTAACTGAACCTGATTTTCACTCTTGG 2272  
Qy 2161 GTGCTACTGTGTTTGTGATCTCTGAGAGCTGGGTCTGGAATTAACCACTACTCCAGT 2220  
Db 2273 GTGCTACTGTGTTTGTGATCTCTGAGAGCTGGGTCTGGAATTAACCACTACTCCAGT 2332  
Qy 2221 GAGCTCCAGGCTTCCAGCACTTTTCAATTTTGTGAATCTTCCCTACTGCTATACGA 2280  
Db 2333 GAGCTCCAGGCTTCCAGCACTTTTCAATTTTGTGAATCTTCCCTACTGCTATACGA 2392  
Qy 2281 GGTAAGAAATTTGCTTGAAGATTAATTAATTAATTAATTAATTAATTAATTAATTA 2340  
Db 2393 GGTAAGAAATTTGCTTGAAGATTAATTAATTAATTAATTAATTAATTAATTAATTA 2452  
Qy 2341 AAGTAAATCAATTAAGAAAGTGAACAAATTTGATATTTCAATGACTTCAAGTGAATTA 2400  
Db 2453 AAGTAAATCAATTAAGAAAGTGAACAAATTTGATATTTCAATGACTTCAAGTGAATTA 2512  
Qy 2401 GCCACAGGCAACAGAGAACCTTCTGTGTTCCAGTGAAGATGAGGCACTGTTCTTTT 2460  
Db 2513 GCCACAGGCAACAGAGAACCTTCTGTGTTCCAGTGAAGATGAGGCACTGTTCTTTT 2572  
Qy 2461 CCCATCAGGCAACATCTGAGAGAAATTTCTATCAAGTCAAGCTCTTTCACCCACT 2520  
Db 2573 CCCATCAGGCAACATCTGAGAGAAATTTCTATCAAGTCAAGCTCTTTCACCCACT 2632

**THIS PAGE BLANK (USPTO)**

2521 GCTTGTGATGCTGTGACCGAGATGATTTTAAAGCTGAGAGAAATGAAAAATCATAT 2580  
 2633 GCTTGTGATGCTGTGACCGAGATGATTTTAAAGCTGAGAGAAATGAAAAATCATAT 2692  
 2581 TCACATATCATCTTATTTAGACTGACGCAATAGAGCTACAGAGTACCTGAAAATTGG 2640  
 2693 TCACATATCATCTTATTTAGACTGACGCAATAGAGCTACAGAGTACCTGAAAATTGG 2752  
 2641 AGTTTCTCATTTCTCTCTAATACAGTGAATGAGTGAAGAAAGTTCAGATCACTGCAAT 2700  
 2753 AGTTTCTCATTTCTCTCTAATACAGTGAATGAGTGAAGAAAGTTCAGATCACTGCAAT 2812  
 2701 GAGAGATGTTCTGTGCTTCTTCCATGATGAGTGAAGTGAAGTGAAGTGAAGTGAAG 2760  
 2813 GAGAGATGTTCTGTGCTTCTTCCATGATGAGTGAAGTGAAGTGAAGTGAAGTGAAG 2872  
 2761 TGTGTGAAACAGAAACATGATTAATTTTGTCTCCAAATATTATTAATTTGATTTGACT 2820  
 2873 TGTGTGAAACAGAAACATGATTAATTTTGTCTCCAAATATTATTAATTTGATTTGACT 2932  
 2821 AAAAAAACAACCTGACAGATTAATTTGAAAAGAAAGTCTTTGATTAATGAGCAAGGT 2880  
 2933 AAAAAAACAACCTGACAGATTAATTTGAAAAGAAAGTCTTTGATTAATGAGCAAGGT 2992  
 2881 TACAGAGAACTTCTCTATGAGAGAAAGTGTCTTTGAGTGTCTTTGAGTGTCTTTGAG 2940  
 2993 TACAGAGAACTTCTCTATGAGAGAAAGTGTCTTTGAGTGTCTTTGAGTGTCTTTGAG 3052  
 2941 GACCTTCTGAGAGCACTTGTGTGAGTGTCTTTGAGTGTCTTTGAGTGTCTTTGAG 3000  
 3053 GACCTTCTGAGAGCACTTGTGTGAGTGTCTTTGAGTGTCTTTGAGTGTCTTTGAG 3112  
 3001 CCTTACATGATTTGATGAGAAATGTTTACAGAAACATACCTTGTGATTAAGCAAT 3060  
 3113 CCTTACATGATTTGATGAGAAATGTTTACAGAAACATACCTTGTGATTAAGCAAT 3172  
 3061 CAGAAATCCAAAGGTGAAATTTTGGATCCAGAAAGTGTCTTTGAGTGTCTTTGAG 3120  
 3173 CAGAAATCCAAAGGTGAAATTTTGGATCCAGAAAGTGTCTTTGAGTGTCTTTGAG 3232  
 3121 GGCATTAATAAGTCAAGTAACTTACAGCTTATTTGAACTTCTCTGAGGATTAAGA 3180  
 3233 GGCATTAATAAGTCAAGTAACTTACAGCTTATTTGAACTTCTCTGAGGATTAAGA 3292  
 3181 AAGTATCAGCTTACATGATGAGTGAAGTGTCTTTGAGTGTCTTTGAGTGTCTTTGAG 3240  
 3293 AAGTATCAGCTTACATGATGAGTGAAGTGTCTTTGAGTGTCTTTGAGTGTCTTTGAG 3352  
 3241 AAGAGATTTTACAGAAATTAATCTAGCTTATTAATTTGAGTGTCTTTGAGTGTCTTTGAG 3300  
 3353 AAGAGATTTTACAGAAATTAATCTAGCTTATTAATTTGAGTGTCTTTGAGTGTCTTTGAG 3412  
 3301 AGTCTTAAAGCAGAAAGCTTGTGATTAATGCTGATTTGAGAGCAGAAACAGAAAGTGTGC 3360  
 3413 AGTCTTAAAGCAGAAAGCTTGTGATTAATGCTGATTTGAGAGCAGAAACAGAAAGTGTGC 3472  
 3361 ATGCAATTTGTGAGTGTCTTACAGAGTCCAACTTTGAGTGTCTTGTGAGAGCAGCTCTCT 3420  
 3473 ATGCAATTTGTGAGTGTCTTACAGAGTCCAACTTTGAGTGTCTTGTGAGAGCAGCTCTCT 3532  
 3421 GATATTGAGTGTGAGCTTATGAGTGTCTTACAGTGTCTTCACTTTGAAATTTGAGTGTCTT 3480  
 3533 GATATTGAGTGTGAGCTTATGAGTGTCTTACAGTGTCTTCACTTTGAAATTTGAGTGTCTT 3592

3601 GAAAGACAAATATTCAGAGTACCGGTGACGGGGCTTACCTACCAAGTCTT----- 3651  
 3713 GAAAGACAAATATTCAGAGTACCGGTGACGGGGCTTACCTACCAAGTCTT----- 3772  
 3652 -----CTTGTGTGTGAGAGCA 3669  
 3773 CTGATTGACACACACACACCGCTTACTCTTACAGACAGAGAGTGTGTGTGAGAGCA 3832  
 3670 ATGAGATTAATTAATTTCCGCAATGTTTGTGATTTGATTTGATTTGATTTGATTTGAT 3729  
 3833 ATGAGATTAATTAATTTCCGCAATGTTTGTGATTTGATTTGATTTGATTTGATTTGAT 3892  
 3730 TATTAATGAGAGGCTTGTGAGTCTTCTGAGAGAGCAAGTCTTCAATTAATTAAGAGCC 3789  
 3893 TATTAATGAGAGGCTTGTGAGTCTTCTGAGAGAGCAAGTCTTCAATTAATTAAGAGCC 3952  
 3790 TTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGAT 3849  
 3953 TTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGAT 4012  
 3850 GTGTGTACAAAGCTTTTGGGCGCGGTGAGAGTGTGATGCTTTATGAGAGTAACTTA 3909  
 4013 GTGTGTACAAAGCTTTTGGGCGCGGTGAGAGTGTGATGCTTTATGAGAGTAACTTA 4072  
 3910 TTAAGTGTGCTTTATGAGTGTGCTTCAAGAGCAATTTCTGAGAGAGACAGTGAAGAGT 3969  
 4073 TTAAGTGTGCTTTATGAGTGTGCTTCAAGAGCAATTTCTGAGAGAGACAGTGAAGAGT 4132  
 3970 GAAATATGATCATGAGAAACTCAACTCTATTAGATTTGATTTGATTTGATTTGATTTGAT 4029  
 4133 GAAATATGATCATGAGAAACTCAACTCTATTAGATTTGATTTGATTTGATTTGATTTGAT 4192  
 4030 GTTAATATTTCTGTGTGAGAAACTTTAAAGTTTCAATTAATCCCAAGATGCTTCACTGTC 4089  
 4193 GTTAATATTTCTGTGTGAGAAACTTTAAAGTTTCAATTAATCCCAAGATGCTTCACTGTC 4252  
 4090 ATAGTGTATTAATTAAGACCAAGAGACAGAGCGGTGAGAAAGTGAACATCTGAAGTGAAG 4149  
 4253 ATAGTGTATTAATTAAGACCAAGAGACAGAGCGGTGAGAAAGTGAACATCTGAAGTGAAG 4312  
 4150 CTGTCTCTGTGAGACTTTTGCAGTGTGCAAGGCTCCGTCTTGTGAGAGTGAAGT 4209  
 4313 CTGTCTCTGTGAGACTTTTGCAGTGTGCAAGGCTCCGTCTTGTGAGAGTGAAGT 4372  
 4210 TCAAGCTCCATCATCACTCTTCACTGATTTTATTTTGTGTTCAAGCTTGTGATTTT 4269  
 4373 TCAAGCTCCATCATCACTCTTCACTGATTTTATTTTGTGTTCAAGCTTGTGATTTT 4432  
 4270 ATGGAACCTTGTGCTGGA 4287  
 4433 ATGGAACCTTGTGCTGGA 4450

RESULT 3  
 AX534938  
 LOCUS AX534938 5895 bp DNA linear PAT 22-NOV-2002  
 DEFINITION Sequence 5 from Patent WO02070696.  
 ACCESSION AX534938  
 VERSION AX534938.1 GI:25261482  
 KEYWORDS  
 SOURCE  
 ORGANISM  
 Homo sapiens (human)  
 Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Platyrrhini; Catarrhini; Hominoidea; Homo.  
 REFERENCE  
 1 Schuch, A. and Sutherland, R.D.  
 Cdio nucleic acid molecules, polypeptides and methods of use  
 Patent: WO 02070696-A 5 12-SEP-2002;  
 Schmidt, Andre (ex) Sutherland, Robert D. (CA)  
 Location/Qualifiers  
 FEATURES  
 source  
 1..5895  
 /organism="Homo sapiens"  
 /mol\_type="genomic DNA"

**THIS PAGE BLANK (USPTO)**

98.5%	Score 4221.4;	DB 6;	Length 4335;
98.8%	Pred. NO. 0;		
0;	Mismatches	1;	Indels 51; Gaps 1

[illegible]

**THIS PAGE BLANK (USPTO)**



1081	AGCCATCTCTCACTTCAAGCCAGCTGTAAGGTAATCTGCTGATGGAACCACTG	1140
1141	ACTCTTGAAGAAAGAAATTAATCTATCTTAACAGTGACACAGAACTATCTAG	1200
1141	ACTCTTGAAGAAAGAAATTAATCTATCTTAACAGTGACACAGAACTATCTAG	1200
1201	TACTGAGCGGATCTAACAGGTGAATCAAGAAATGGAAGCTGTCGAAATTAAT	1260
1201	TACTGAGCGGATCTAACAGGTGAATCAAGAAATGGAAGCTGTCGAAATTAAT	1260
1261	ACTGCCCCCAAACTGGAATTTTAAGATTGAATTCCTGAGAGATTCAAGTGA	1320
1261	ACTGCCCCCAAACTGGAATTTTAAGATTGAATTCCTGAGAGATTCAAGTGA	1320
1321	CTACAGTTGAAGGCTTATTTCTTGTGTAAGAAAGTCAATTCCTGAGAGATT	1380
1321	CTACAGTTGAAGGCTTATTTCTTGTGTAAGAAAGTCAATTCCTGAGAGATT	1380
1381	AACTCTCTAGTAAGACATATCACTAACTAAACAGAGATGAATTAATTAAGT	1440
1381	AACTCTCTAGTAAGACATATCACTAACTAAACAGAGATGAATTAATTAAGT	1440
1441	TGCGCTTTGAGTTGGTGTAGTGCAAAAGATTGAAGAGATTGAATTAAGT	1500
1441	TGCGCTTTGAGTTGGTGTAGTGCAAAAGATTGAAGAGATTGAATTAAGT	1500
1501	GTAATCCAGGAGACAGTTGGTGTAGTGCAAAAGATTGAAGAGATTGAAT	1560
1501	GTAATCCAGGAGACAGTTGGTGTAGTGCAAAAGATTGAAGAGATTGAAT	1560
1561	CCAGAAATTTCTTGAATCTCAAAAGCTGTGTATTTGTAATTAATTAAGT	1620
1561	CCAGAAATTTCTTGAATCTCAAAAGCTGTGTATTTGTAATTAATTAAGT	1620
1621	GAATTAATTAAGTGTCTTAAATTTCTGTTCACTGTTGTTTAAATTAAGT	1680
1621	GAATTAATTAAGTGTCTTAAATTTCTGTTCACTGTTGTTTAAATTAAGT	1680
1681	CTATATTGAGAGTAAGTGAAGTGAATGCTGTAATTTGTAATTAAGT	1740
1681	CTATATTGAGAGTAAGTGAAGTGAATGCTGTAATTTGTAATTAAGT	1740
1741	ACACGCTGACTCCATAGTTGGATTTGAGCTGTTGCAAAAGTGAATGTAAT	1800
1741	ACACGCTGACTCCATAGTTGGATTTGAGCTGTTGCAAAAGTGAATGTAAT	1800
1801	GCTCTAATGATTAATTAAGTGAATTTGCTCAATGTTGAATTTTAACACAGA	1860
1801	GCTCTAATGATTAATTAAGTGAATTTGCTCAATGTTGAATTTTAACACAGA	1860
1861	TATATTAGGATCTTCAATGATTTCTTTGAGCTTTCAGAAATGGAATCTG	1920
1861	TATATTAGGATCTTCAATGATTTCTTTGAGCTTTCAGAAATGGAATCTG	1920
1921	TTGACAGATGCAAACTCAAGAGATTAATGATGTTTATGACAAATGAGAAAT	1980
1921	TTGACAGATGCAAACTCAAGAGATTAATGATGTTTATGACAAATGAGAAAT	1980
1981	GCTGAGAGTTTATGAGAGAAATGAAGACATTTGTAATTTCAATGCTTTTG	2040
1981	GCTGAGAGTTTATGAGAGAAATGAAGACATTTGTAATTTCAATGCTTTTG	2040
2041	GCTGAGAGTTTATGAGAGAAATGAAGACATTTGTAATTTCAATGCTTTTG	2100
2041	GCTGAGAGTTTATGAGAGAAATGAAGACATTTGTAATTTCAATGCTTTTG	2100
2101	ATGAGTTTATGAGAGAAATGAAGACATTTGTAATTTCAATGCTTTTG	2160
2101	ATGAGTTTATGAGAGAAATGAAGACATTTGTAATTTCAATGCTTTTG	2160
2161	GTGCTACTGTTTGTGATCTCTGAGACCTGGCTTGAATTAACATCTCAGT	2220
2161	GTGCTACTGTTTGTGATCTCTGAGACCTGGCTTGAATTAACATCTCAGT	2220
2221	GAGTCCAGGCTTCCAACTATTTTCTTATTTTGAATCTTCCCTATCTGTTAT	2280
2221	GAGTCCAGGCTTCCAACTATTTTCTTATTTTGAATCTTCCCTATCTGTTAT	2280
2281	GCTGAGAAATTTGCTTGGAAATTAATTAATTTGTAATTTGTAATTTGTAAT	2340
2281	GCTGAGAAATTTGCTTGGAAATTAATTAATTTGTAATTTGTAATTTGTAAT	2340
2341	AAAGTAATCATTTGAGAAATGAAGAAATTTGTAATTTGTAATTTGTAAT	2400
2341	AAAGTAATCATTTGAGAAATGAAGAAATTTGTAATTTGTAATTTGTAAT	2400
2401	GCACAGGCCACAGACAGCTTCTGTTTCCAGTGAAGATGGGCACTGTTCT	2460
2401	GCACAGGCCACAGACAGCTTCTGTTTCCAGTGAAGATGGGCACTGTTCT	2460
2461	CCCATCAGGCCACAGACAGCTTCTGTTTCCAGTGAAGATGGGCACTGTT	2520
2461	CCCATCAGGCCACAGACAGCTTCTGTTTCCAGTGAAGATGGGCACTGTT	2520
2521	GCTTGTATCTGTCAACAGATGATTTTAAAGGCTGAAGAAATGAATTAAT	2580
2521	GCTTGTATCTGTCAACAGATGATTTTAAAGGCTGAAGAAATGAATTAAT	2580
2581	TCCATATCCATCTTATTAAGCTTGAAGTGAAGTGAAGTGAAGTGAAGT	2640
2581	TCCATATCCATCTTATTAAGCTTGAAGTGAAGTGAAGTGAAGTGAAGT	2640
2641	AGTTTCTATTTCTCTTAAATTAAGTGAAGTGAAGTGAAGTGAAGTGAAG	2700
2641	AGTTTCTATTTCTCTTAAATTAAGTGAAGTGAAGTGAAGTGAAGTGAAG	2700
2701	GAGATGTTCTGTTCTTCAATCAATGAGTGAAGTGAAGTGAAGTGAAGT	2760
2701	GAGATGTTCTGTTCTTCAATCAATGAGTGAAGTGAAGTGAAGTGAAGT	2760
2761	TGAGGAAACAGAAATTAATTTGCTCAATTAATTTGTAATTTGTAATTT	2820
2761	TGAGGAAACAGAAATTAATTTGCTCAATTAATTTGTAATTTGTAATTT	2820
2821	AAAAAGAAACAACTGACATTAATTTGAAAGAAAGCTCTTCAATTTATG	2880
2821	AAAAAGAAACAACTGACATTAATTTGAAAGAAAGCTCTTCAATTTATG	2880
2881	TACAGAGAACTTCTTATCAAGAGAAAGTGTCTTCAATTTATGAGCAAGT	2940
2881	TACAGAGAACTTCTTATCAAGAGAAAGTGTCTTCAATTTATGAGCAAGT	2940
2941	GACCTTCTGAGACACTGTTGTGACGTTTGTATTAAGATTTTCTTGAAG	3000
2941	GACCTTCTGAGACACTGTTGTGACGTTTGTATTAAGATTTTCTTGAAG	3000
3001	CCTTACATAGATTTGATGCAATTTGTAATTTGTAATTTGTAATTTGTAAT	3060
3001	CCTTACATAGATTTGATGCAATTTGTAATTTGTAATTTGTAATTTGTAAT	3060
3061	CAGAAATCCAACTGTAATTTGAGATCCAGAGAGATTTCAATGAGCTTCA	3120
3061	CAGAAATCCAACTGTAATTTGAGATCCAGAGAGATTTCAATGAGCTTCA	3120
3121	GGCAATTAAGTCAATTAACATTTGTAATTTGTAATTTGTAATTTGTAAT	3180
3121	GGCAATTAAGTCAATTAACATTTGTAATTTGTAATTTGTAATTTGTAAT	3180
3181	AAATTAAGTCAATTAACATTTGTAATTTGTAATTTGTAATTTGTAATTT	3240
3181	AAATTAAGTCAATTAACATTTGTAATTTGTAATTTGTAATTTGTAATTT	3240
3241	AGAGAAATTTGAGAAATTAATTTGTAATTTGTAATTTGTAATTTGTAAT	3300
3241	AGAGAAATTTGAGAAATTAATTTGTAATTTGTAATTTGTAATTTGTAAT	3300

**THIS PAGE BLANK (USPTO)**



**THIS PAGE BLANK (USPTO)**



Creation date: 02-19-2004  
Indexing Officer: AGOMEZ - ALFREDO GOMEZ, JR.  
Team: OIPEBackFileIndexing  
Dossier: 09915515

Legal Date: 01-05-2004

No.	Doccode	Number of pages
1	SRNT	64

Total number of pages: 64

Remarks:

Order of re-scan issued on .....

